



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 120365

TO: Vanessa L Ford  
Location: rem/3b25/3c18  
Art Unit: 1645  
Tuesday, April 27, 2004

Case Serial Number: 10620049

From: Barb O'Bryen  
Location: Biotech-Chem Library  
Remsen 1A69  
Phone: 571-272-2518 *BOB*

barbara.obryen@uspto.gov

### Search Notes

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2004, 15:46:42 ; Search time 2545 Seconds

(without alignments)  
8694.647 Million cell updates/sec

Title: US-10-620-049-24

Perfect score: 741

Sequence: 1 Gattacagacacacagac.....cctcagtcacgcgtctctctcg 741

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pin:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_pig:\*

27: em\_gss\_vrt:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	293	39.5	616	10	BF136295
2	293	39.5	941	10	BF138189
3	292.4	39.5	669	10	BF123422
4	282	38.1	462	14	CA578969

5	282	38.1	481	14	CA575938
6	281.8	38.0	955	13	BQ947728
7	281.2	37.9	827	12	BI152061
8	281.2	37.9	1459	13	BQ918407
9	280.4	37.8	774	10	BF123744
10	277.2	37.4	891	10	BF138460
11	272.4	36.8	896	12	BI149320
12	270.2	36.5	589	11	AK007163
13	269	36.3	708	12	BI103114
14	268.2	36.2	940	10	BF169023
15	267.8	36.1	417	14	CA577519
16	267.4	36.1	429	14	CA579143
17	267.4	36.1	475	14	CA578534
18	266	35.9	830	10	BF144339
19	265.8	35.9	772	10	BF581992
20	262	35.4	830	12	BI455041
21	260.8	35.2	933	13	BQ929209
22	259.2	35.0	690	12	BG968682
23	257.2	34.7	626	10	BF582283
24	257.2	34.7	725	10	BF580940
25	257	34.7	538	14	CA578968
26	256.4	34.6	724	12	BG962137
27	256	34.5	966	10	BF578083
28	255.4	34.5	979	10	BF180947
29	254.8	34.4	506	14	CA578474
30	254.8	34.4	506	14	CA578686
31	254.2	34.3	465	14	CA578116
32	254	34.3	958	13	BUS23796
33	253	34.1	735	12	BG966898
34	253	34.1	1919	10	BF180034
35	252.6	34.1	698	10	BF584024
36	251.6	34.0	1576	11	AK007918
37	250.8	33.8	794	12	BI150371
38	250.6	33.8	739	10	BE284158
39	250.6	33.8	931	13	BUS23606
40	250.2	33.8	439	10	BE136816
41	248.4	33.5	596	10	BE310080
42	247.4	33.4	863	13	BQ952037
43	247.4	33.4	916	10	BF163514
44	247	33.3	469	14	CA574907
45	246.6	33.3	473	14	CA580167

#### ALIGNMENTS

RESULT 1  
BF136295  
LOCUS 601781413F1 NCI\_CGAP\_Lu30 Mus musculus cdna clone IMAGE:4009443 5',  
DEFINITION 616 bp mRNA linear EST 24-OCT-2000  
mRNA sequence.  
ACCESSION BF136295  
VERSION BF136295.1 GI:10975335  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 616)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Gilbert Smith, Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLNM9245 row: j column: 04  
High quality sequence stop: 613.

CA576938 K0702C06-  
BQ947728 AGENCOURT  
BI152061 602916512  
BQ918407 AGENCOURT  
BF123744 601760491  
BF138460 601782916  
BI149320 602848859  
AK007163 Mus muscu  
BI103114 602899345  
BF169023 601775295  
CA577519 K0710F01-  
CA579143 K0733D01-  
CA578534 K0735B11-  
BF144339 601787158  
BF581992 602099448  
BI455041 603173343  
BQ929209 AGENCOURT  
BG968682 602836513  
BF582283 602101109  
BF580940 602100636  
CA578968 K0731A04-  
BG962137 602836902  
BF578083 602094759  
BF180947 601807187  
CA578474 K0724C11-  
CA578686 K0727C05-  
CA578116 K0718G11-  
BUS23796 AGENCOURT  
BG966898 602834104  
BF180034 601806383  
BF584024 602086264  
AK007918 Mus muscu  
BI150371 602915205  
BE284158 601099428  
BUS23606 AGENCOURT  
BE136816 U956501.V  
BE310080 601091549  
BQ952037 AGENCOURT  
BF163514 601771691  
CA574907 K0622H02-  
CA580167 K0747G08-

## FEATURES

source  
Location/Qualifiers  
1. .616  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="Czech II"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4009443"  
/issue\_type="tumor, metastatic to mammary"  
/lab\_host="DH10B"  
/clone\_lib="NCI\_CGAP\_Lu30"  
/note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; transgenic model WNT-1, expression driven by  
WNTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo  
dT. Library constructed by Life Technologies.  
Investigator providing samples: Gilbert Smith, NIH"

## ORIGIN

Query Match 39.5%; Score 293; DB 10; Length 616;  
Best Local Similarity 87.9%; Pred. No. 1.4e-58;  
Matches 333; Conservative 0; Mismatches 40; Indels 6; Gaps 1;  
QY 368 CCGGTGGTGGATCGAGTCCAACTGCAACAGTCTGGACCTGAGCTGGTGAAGCCTG 427  
Db 49 CTGAAGTGTCCACTCCAGTTCAGCTGCAGCAGTCTGGACCTGAGCTGGTGAAGCCTG 108  
QY 428 GGGCTCAGTGAAGATTCTCGAAAGATTCTGGTACGCAATCAATAGCTCTTGGATGA 487  
Db 109 GGGCTCAGTGAAGATTCTCGAAAGATTCTGGTACGCAATCAATAGCTCTTGGATGA 168  
QY 488 ACTGGGTGAGCAGAGCCCTGGACAGGCTCTTCAGTGGATTCGACGGAATTCCTCTGGAG 547  
Db 169 ACTGGGTGAGCAGAGCCCTGGACAGGCTCTTCAGTGGATTCGACGGAATTCCTCTGGAG 228  
QY 548 ATGGAGATTCTTAATCAATGGGAATTCGAGGGCAAGGCCATCTAGCAGCAGCAAAAT 607  
Db 229 ATGGAGATTCTTAATCAATGGGAATTCGAGGGCAAGGCCATCTAGCAGCAGCAAAAT 288  
QY 608 CTTCCAGCAGCCTACATGCTAGCTGAGCAGCTGCTGCTGGATCTGGCTGCTATT 667  
Db 289 CTTCCAGCAGCCTACATGCTAGCTGAGCAGCTGCTGCTGGATCTGGCTGCTATT 348  
QY 668 TCTGTGCAAGATCGGGTGTCTAGCTT-----ATGCTATGACTACTGGGTCAAGGAA 721  
Db 349 TCTGTGCAAGATTTAACTACTATGCTTCTACAATGCTATGGACTACTGGGTCAAGGAA 408  
QY 722 CTTCACTACCGTCTCTCTC 740  
Db 409 CTTCACTACCGTCTCTCTC 427

RESULT 2  
BF138189  
LOCUS 601781314F1 NCI\_CGAP\_Lu30 941 bp mRNA linear EST 24-OCT-2000  
DEFINITION 601781314F1 NCI\_CGAP\_Lu30 Mus musculus cDNA clone IMAGE:4009467 5',  
mRNA sequence.  
ACCESSION BF138189  
VERSION BF138189.1 GI:10977229  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 941)  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov  
Plate: LLAM9245 row: k column: 04  
High quality sequence stop: 637.

## FEATURES

source  
Location/Qualifiers  
1. .941  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="Czech II"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4009467"  
/issue\_type="tumor, metastatic to mammary"  
/lab\_host="DH10B"  
/clone\_lib="NCI\_CGAP\_Lu30"  
/note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; transgenic model WNT-1, expression driven by  
WNTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo  
dT. Library constructed by Life Technologies.  
Investigator providing samples: Gilbert Smith, NIH"

## ORIGIN

Query Match 39.5%; Score 293; DB 10; Length 941;  
Best Local Similarity 87.9%; Pred. No. 1.7e-58;  
Matches 333; Conservative 0; Mismatches 40; Indels 6; Gaps 1;  
QY 368 CCGGTGGTGGATCGAGTCCAACTGCAACAGTCTGGACCTGAGCTGGTGAAGCCTG 427  
Db 48 CTGAAGTGTCCACTCCAGTTCAGCTGCAGCAGTCTGGACCTGAGCTGGTGAAGCCTG 107  
QY 428 GGGCTCAGTGAAGATTCTCGAAAGATTCTGGTACGCAATCAATAGCTCTTGGATGA 487  
Db 108 GGGCTCAGTGAAGATTCTCGAAAGATTCTGGTACGCAATCAATAGCTCTTGGATGA 167  
QY 488 ACTGGGTGAGCAGAGCCCTGGACAGGCTCTTCAGTGGATTCGACGGAATTCCTCTGGAG 547  
Db 168 ACTGGGTGAGCAGAGCCCTGGACAGGCTCTTCAGTGGATTCGACGGAATTCCTCTGGAG 227  
QY 548 ATGGAGATTCTTAATCAATGGGAATTCGAGGGCAAGGCCATCTAGCAGCAGCAAAAT 607  
Db 228 ATGGAGATTCTTAATCAATGGGAATTCGAGGGCAAGGCCATCTAGCAGCAGCAAAAT 287  
QY 608 CTTCCAGCAGCCTACATGCTAGCTGAGCAGCTGCTGCTGGATCTGGCTGCTATT 667  
Db 288 CTTCCAGCAGCCTACATGCTAGCTGAGCAGCTGCTGCTGGATCTGGCTGCTATT 347  
QY 668 TCTGTGCAAGATCGGGTGTCTAGCTT-----ATGCTATGACTACTGGGTCAAGGAA 721  
Db 348 TCTGTGCAAGATTTAACTACTATGCTTCTACAATGCTATGGACTACTGGGTCAAGGAA 407  
QY 722 CTTCACTACCGTCTCTCTC 740  
Db 408 CTTCACTACCGTCTCTCTC 426

## RESULT 3

BF123422  
LOCUS 601760623F1 NCI\_CGAP\_Mam5 669 bp mRNA linear EST 24-OCT-2000  
DEFINITION 601760623F1 NCI\_CGAP\_Mam5 Mus musculus cDNA clone IMAGE:4023749 5',  
mRNA sequence.  
ACCESSION BF123422  
VERSION BF123422.1 GI:10962462  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 669)  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys

http://image.llnl.gov  
Plate: LLAM9282 row: n column: 06  
High quality sequence stop: 649.

/clone="Lib-NC1-CGAL" name="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dr. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Robin Humphreys, NIH"

(EHII-7/C-KIC-7/SCA-1-1) CDNA

[illegible]

608

[illegible]

Db 316 CCTCAGCAGCAGCTACATGCACTCAGCAGCCTGACATCTGAGACTCTGGGCTCTACT 375  
 QY 668 TCTGTGCAAGATCGGGGTGCTACGTTATGCTATGCACTACTGGGTCAGCAAGCACTCAG 727  
 Db 376 TCTGTGCAAGATCAGCTAGCTAGCTATAGGTTGCTTACTGGGCGCCAGGAGCTCTCG 435  
 QY 728 TCACCGTCTC 737  
 Db 436 TCACGTCTC 445

RESULT 5  
 CAS76938  
 LOCUS  
 DEFINITION  
 K0702C06-5N NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1-)  
 cDNA Library Mus musculus cDNA clone KIA:K0702C06  
 IMAGE:30073757 5', mRNA sequence.

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Mus musculus (house mouse)  
 EST.  
 CAS76938.1 GI:25125329  
 Mus musculus (house mouse)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 1 (bases 1 to 481)  
 Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Luo, A., Carter, M.G.,  
 Aiba, K., Taub, D., Longo, D.L., Keller, J., and Ko, M.S.H.  
 Systematic Analyses of NIA Mouse Hematopoietic Stem Cell  
 (Lin-/c-Kit-/Sca-1-) cDNA Library (Long)  
 Unpublished (2001)  
 Contact: Dawood B. Dudekula  
 Laboratory of Genetics  
 National Institute on Aging/National Institutes of Health  
 333 Caswell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
 Email: cdna@lsun.grc.nia.nih.gov  
 Plate: K0702 row: C column: 06  
 Seq primer: M13 Reverse  
 High quality sequence stop: 481  
 POLYA=No.

FEATURES  
 source  
 1. .481  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6Ncr"  
 /db\_xref="taxon:10090"  
 /clone="NIA:K0702C06 IMAGE:30073757"  
 /tissue\_type="Hematopoietic Stem Cell"  
 (Lin-/c-Kit-/Sca-1-)  
 /dev\_stage="Age approx. 10 weeks old"  
 /lab\_host="PH108"  
 /clone\_lib="NIA Mouse Hematopoietic Stem Cell"  
 (Lin-/c-Kit-/Sca-1-) cDNA Library (Long)  
 /note="Vector: pSPORT1 (Invitrogen); Site\_1: SalI; Site\_2:  
 NotI; Mouse cDNA project by the Laboratory of Genetics,  
 National Institute on Aging (NIA), Intramural Research  
 Program, NIH (http://lsun-grc.nia.nih.gov/cDNA). This is  
 a long-transcript enriched cDNA library (Ref. Genome Res.  
 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were  
 obtained from Drs. Dennis Taub, Dan Longo (National Cancer  
 Institute on Aging, USA), Jonathan Keller (National Cancer  
 Institute, USA). Double-stranded cDNAs were synthesized  
 with an Oligo(dT) primer [Invitrogen]:  
 5'-pGACTAGTCTAGTCGCGCGCCCTTTT-3' from  
 0.9 ug of total RNA, treated with T4 DNA polymerase, and  
 purified by ethanol-precipitation. The cDNAs were ligated  
 to Loner-linker LL-Sal4, purified by phenol/chloroform, and  
 separated from free linkers by Centricon 100. Then, the  
 cDNAs were amplified by long-range high fidelity PCR using  
 Ex Taq polymerase (Takara) with a primer Sal4-S. The  
 products were purified by phenol/chloroform and Centricon  
 100. The cDNAs were digested with SalI and NotI enzymes  
 and cloned into SalI/NotI site of pSPORT1 plasmid vector.

The DH10B *E. coli* host was transformed with the ligation  
 mixture by the standard chemical method. The average  
 insert size is about 2.1 kb. The library was constructed  
 by Yulan Piao (NIA)."

ORIGIN  
 Query Match 38.1%; Score 282; DB 14; Length 481;  
 Best Local Similarity 85.1%; Pred. No. 5.2e-56;  
 Matches 315; Conservative 0; Mismatches 55; Indels 0; Gaps 0;  
 QY 368 CCGTGTGTGGATCCGAGGTCCAACTGCAACAGTCTGACCTGAGTGGTGAACCTG 427  
 Db 76 CTGAAGGTGTCCACTCCCAAGGTTTCAGCTGCAGCAGTCTGACCTGAGTGGTGAACCTG 135  
 QY 428 GGGCTCTAGTCAAGATTCTCTGCAAGATTCTGCTACGCAATCAATAGCTCTTGGATGA 487  
 Db 136 GGGCTCTAGTCAAGATTCTCTGCAAGCTTCTGCTACGCAATCAATAGCTCTTGGATGA 195  
 QY 488 ACTGGTGAAGCAGAGCGCTGGAACAGGCTCTTGAGTGGATTGACGGAATTCCTGGAG 547  
 Db 196 ACTGGTGAAGCAGAGCGCTGGAAGAGGCTCTTGAGTGGATTGACGGAATTCCTGGAG 255  
 QY 548 ATGGAGATTCTAACTACAAATGGGAAATTCGAGGCAAGGCCACTACTGACAGCAGACAAAT 607  
 Db 256 ATGGAGATTCTAACTACAAATGGGAAATTCGAGGCAAGGCCACTACTGACAGCAGACAAAT 315  
 QY 608 CTCCTCAGCAGCCTACATGCAAGTCTGAGCAGCTGAGCTCTGAGGACTCTGCGGTCTATT 667  
 Db 316 CTCCTCAGCAGCCTACATGCAAGTCTGAGCAGCTGAGCTCTGAGGACTCTGCGGTCTATT 375  
 QY 668 TCTGTGCAAGATCGGGTTCGAGTCTATGCTATGAGTCTGCTTACTGGGTCAGGAACTCTCAG 727  
 Db 376 TCTGTGCAAGATCGGGTTCGAGTCTATGAGTCTGCTTACTGGGTCAGGAACTCTCAG 435  
 QY 728 TCACGTCTC 737  
 Db 436 TCACGTCTC 445

RESULT 6  
 BQ947728  
 LOCUS  
 DEFINITION  
 BQ947728 955 bp mRNA linear EST 21-AUG-2002  
 IMAGE:6394985 5', mRNA sequence.  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 1 (bases 1 to 955)  
 NIH-MGC http://mgc.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: The Cepko Laboratory  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM1388 row: 0 column: 18  
 High quality sequence stop: 712.  
 Location/Qualifiers  
 1. 955  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:6394985"



```

/clone="IMAGE:4023550"
/tissue_type="tumor_gross_tissue"
/dev_stage="7 months"
/lab_post="DH103"
/clone_lib="NCI CGAP Mams"
/note="Organ: mammary; Vector: pcWV-SPORE; Site 1: Sall;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dr.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH."

```

## ORIGIN

Query Match	37.8%; Score 280.4; DB 10; Length 774;
Best Local Similarity	91.9%; Pred. No. 1.5e-55;
Matches 307; Conservative	0; Mismatches 26; Indels 1; Gaps 1;

  

Qy	1	GATATTGAGTGCACAGACTACATCCTCCCTGTCTGCCCTCTCGGAGACAGAGTCACC	60
Db	50	GATATCCAGATGACACAGACTACATCCTCCCTGTCTGCCCTCTCGGAGACAGAGTCACC	108
Qy	61	GTCAAGTTGCAGGGCAAGTCAGGACATTAGGAATTATTAAACTGGTATCAGCAGAAACCA	120
Db	110	ATCAGTTGCAGGGCAAGTCAGGACATTAGCAAATTTTAACTGGTATCAGCAGAAACCA	169
Qy	121	GACGGAACTGTTAAATTCCTGATCTACTACATCAAGATTACTGCCAGAGTGCCCATCA	180
Db	170	GACGGAACTGTTAAACTCCTGATCTATTACACATCAACATTACAACAGGAGTCCCATCA	229
Qy	181	AGGTTCAAGTCGCAGTGGGTCTGAAACAGATTATTCCTCACCATTAACAACCTGGAGCAG	240
Db	230	AGGTTCAAGTCGCAGTGGGTCTGAAACAGATTATTCTCTCACCATTAACAACCTGGAGCAA	289
Qy	241	GAGAGATAATGGCACTTACTTTTTGGCCAACGGGCAATACGGCTCCGTCGGAGAGTTCGGTGGGA	300
Db	290	GAGAGATA-TGCCACTTACTTTTGGCCACAGGTAAGCGCTTCTCTCGAGCGTTCGGTGGGA	348
Qy	301	GGCACCAAGCTGGAAAATAAAAAACGTGGTGGAGGTG	334
Db	349	GGCACCAAGCTGGAAAATAAAAAACGTGGTGGAGGTG	382

## RESULT 10

[illegible]

Email: [cgapbs-t@mail.nih.gov](mailto:cgapbs-t@mail.nih.gov)  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: L1AM9249 row: o column: 04  
High quality sequence stop: 581.

DECLASSIFIED

```

FEATURES
  source
    1. .891
      /organism="Mus musculus"
      /mol_type="mRNA"
      /strain="Czech II"
      high quality sequence study. 201.
      Location/Qualifiers

```



```

/db_xref="taxon:10090"
/clone="IMAGE:4011099"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH108"
/clone_lib="NCI_CGAP_Lu30"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; transgenic model WNT-1, expression driven by
MTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
dT. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"

```

## ORIGIN

```

Query Match      37.4%; Score 277.2; DB 10; Length 891;
Best Local Similarity 88.8%; Pred. No. 9.2e-55;
Matches 324; Conservative 0; Mismatches 34; Indels 7; Gaps 2;

```

```

QY 383 CCGAGGTCAACACTCGACCTGGACCTGGAGCTGGTGAAGCTGGGGCTCAGTGAAGA 442
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 CCAGGTTTCAGCTGCAGCAGCTCGACCTGAGCTGGTGAAGCTGGGGCTCAGTGAAGA 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 443 TTTCTGCGAAAGATTCCTGGCTAGCAGCTATCAATAGCTTTGGATGAAGTGGGTGAAGCAGA 502
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 TTTCTGCGAAGGCTTCTGGCTAGCAGCTATCAGTAGCTCTCTGGATGAAGTGGGTGAAGCAGA 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 503 GGCTCGACAGGGTCTT-GAGTGGATTGACAGGATTTATCTGGAGATGGAGATTCTAAC 561
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 181 GGCTCGAAGGCTCTTGGAGTGGATTGACAGGATTTATCTGGAGATGGAGATTCTAAC 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 562 TACAATGGGAAATTCGAGGCAAGCCATCTACTGACAGCAGACAAATCTCTCCAGCAGCC 621
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 241 TACAATGGGAAATTCGAGGCAAGCCATCTACTGACAGCAGACAAATCTCTCCAGCAGCC 300
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 622 TACATCAGCTCAGAGCTGACCTCTGAGGCTCTGCGGTCTATTCTGTGCAAGATCG 681
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 301 TACATCAGCTCAGAGCTGACCTCTGAGGCTCTGCGGTCTATTCTGTGCAAGATCG 360
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 682 GGGTTGCTAGGTT-----ATGCTATGAGTACTGGGGTCAAGGAACCTCAGTCCCGTC 735
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 361 AACTACTATGGCTTCTACATGCTATGAGTACTGNGGTCAAGGAACCTCAGTCCCGTC 420
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 736 TCCTC 740
| | | | |
Db 421 TCCGC 425

```

## RESULT 11

```

BI149320      896 bp      mRNA      linear      EST 05-JUL-2001
LOCUS      602848859F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:5012343 5',
DEFINITION      mRNA sequence.
ACCESSION      BI149320
VERSION      BI149320.1 GI:14609321
KEYWORDS      EST.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
REFERENCE      1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE      NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M11063 row: e column: 16
High quality sequence stop: 862.
Location/Qualifiers

```

```

1. 896
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Czech II"
/db_xref="taxon:10090"
/clone="IMAGE:5012343"
/tissue_type="spontaneous tumor, metastatic to mammary.
Stem cell origin."
/lab_host="DH108"
/clone_lib="NCI_CGAP_Lu29"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

```

## ORIGIN

```

Query Match      36.8%; Score 272.4; DB 12; Length 896;
Best Local Similarity 84.8%; Pred. No. 1.3e-53;
Matches 319; Conservative 0; Mismatches 51; Indels 6; Gaps 1;

```

```

QY 368 CCGTGTGTGGATCCGAGGTCCAACTGCAACAGTCTGGACCTGGAGCTGGTGAAGCCG 427
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 81 CTGAAGGTGTCCATCCAGGTTTCAGCTGCAGCAGTCTGGACCTGGAGCTGGTGAAGCCG 140
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 428 GGGCTCAGTGAAGATTTCTGCAAGATTTCTGGTACGCAATCAATAGCTCTTGGATGA 487
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 141 GGGCTCAGTGAAGATTTCTGCAAGATTTCTGGTACGCAATCAATAGCTCTTGGATGA 200
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 488 ACTGGGTGAAGCAGAGCCCTGGACAGGTTCTTGGTGGATTGGACGGATTATCTCTGGAG 547
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 201 ACTGGGTGAAGCAGAGCCCTGGAAAGGTTCTTGGTGGATTGGACGGATTATCTCTGGAG 260
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 548 ATGAGATTTCTTAATCAATGGGAATTCGAGGGCAAGGCCATCTACAGCAGACAAAT 607
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 261 ATGAGATTTCTTAATCAATGGGAATTCGAGGGCAAGGCCATCTACAGCAGACAAAT 320
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 608 CTTCCAGCAGAGCTTACATGACGCTCAGAGCCCTGACCTCTGTGGACTCTCGGTCCTATT 667
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 321 CTTCCAGCAGAGCTTACATGACGCTCAGAGCCCTTAACTCTGAGGACTCTCGGTCCTATT 380
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 568 TCTGTGCAAGATCGGGTTGTACGTTA-----TGCATGAGTACTCTGGGTCAAGGAA 721
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 381 TCTGTGCAAGATCGGGTTGTACGTTA-----TGCATGAGTACTCTGGGTCAAGGAA 440
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 722 CTTCCAGTCAAGCTCTC 737
| | | | |
Db 441 CTTCCAGTCAAGCTCTC 456

```

## RESULT 12

```

AK007163      689 bp      mRNA      linear      HTC 20-SEP-2003
LOCUS      Mus musculus adult male testis cDNA, RIKEN full-length enriched
DEFINITION      library, clone:1700110L11 product:immunoglobulin heavy chain 6
              (heavy chain of IgM), full insert sequence.
ACCESSION      AK007163
VERSION      AK007163.1 GI:12840553
KEYWORDS      HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
REFERENCE      1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE      Carninci, P. and Hayashizaki, Y.
JOURNAL      High-efficiency full-length cDNA cloning
MEDLINE      Meth. Enzymol. 303, 19-44 (1999)
PUBMED      99279253
AUTHORS      Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
              Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE      Normalization and subtraction of cap-trapper-selected cDNAs to
              prepare full-length cDNA libraries for rapid discovery of new genes

```



[illegible]

```

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/strain="FVB/N"
/clone="IMAGE:5044690"
/lab_host="NCI CGAP Kid14"
/notes="Organ: kidney; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library."

```

## ORIGIN

```

Query Match      36.3%; Score 269; DB 12; Length 708;
Best Local Similarity 88.0%; Pred. No. 7.2e-53;
Matches 293; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 2 ATATTGAGTACACAGACTACATCTCTCCCTGTCGCTCTCTGGGAGACAGTCCACG 61
Db 73 ATGTCCAGATGACACAGACTACATCTCTCCCTGTCGCTCTCTGGGAGACAGTCCACCA 132
QY 62 TCAGTTCCAGGCAAGTCAGACATATAGGAATATTAACTGGTATCAGCAGAAACCCAG 121
Db 133 TCATTGCAAGTCAGTCAGAACATTCGCCACTTATTGAACTGGTTTCAGCAGAGCCAG 192
QY 122 ACGGAATGTTAAATTCCTGATCTACTACATCAAGATTTACTGCGAGAGTCCCAATCAA 181
Db 193 ATGGAATGTTAAATTCCTGATCTACTACATCAAGATTTACTGCGAGAGTCCCAATCAA 252
QY 182 GATTTCAGTGCAGTGGTCTGGAACAGATTTATCCCTCACCATTACAACTGGACGAG 241
Db 253 GATTTCAGTGCAGTGGTCTGGAACAGATTTATCCCTCACCATTACAACTGGACGAG 312
QY 242 AAGTATTGGCACTTACTTTTGGCAACAGGCGAATAGCCCTCCGTGGAGCGTTCCGTTGGAG 301
Db 313 AGGATATTGGCGCTTATTTTGGCAACAGGCGTATACGCTTCGTTGGAGCGTTCCGTTGGAG 372
QY 302 GCACCAAGCTGGAATTAACAGTGGTGGAGTG 334
Db 373 GCACCTACCTGGAATTAACAGTGGTGGAGTG 405

```

```

RESULT 14
BF169023      940 bp mRNA linear EST 30-OCT-2000
LOCUS        601775295F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:4017214 5',
DEFINITION   mRNA sequence.
VERSION      BF169023
KEYWORDS     EST.
SOURCE       Mus musculus (house mouse)
ORGANISM     Mus musculus
REFERENCE    1 (bases 1 to 940)
AUTHORS     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE       Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
JOURNAL     NIH-MGC http://mgi.nci.nih.gov/.
COMMENT     National Institutes of Health, Mammalian Gene Collection (MGC)
            Contact: Robert Strausberg, Ph.D.
            Email: cgaps-remail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LIA9265 row: m column: 23
            High quality sequence stop: 581.
            Location/Qualifiers
            1..940
            /organism="Mus musculus"
            /mol_type="mRNA"

```

```

/strain="Czech II"
/db_xref="taxon:10090"
/clone="IMAGE:4017214"
/tissue_type="spontaneous tumor, metastatic to mammary.
Stem cell origin."
/lab_host="NCI CGAP Lu29"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

```

## ORIGIN

```

Query Match      36.2%; Score 268.2; DB 10; Length 940;
Best Local Similarity 86.1%; Pred. No. 1.3e-52;
Matches 321; Conservative 0; Mismatches 48; Indels 4; Gaps 2;

QY 368 CGGTGGTGGTGGATCCAGGTCCAACTGCAACAGTCTGGACCTGAGCTGGTGAAGCTG 427
Db 52 CTGAAGGTGTCCAAATCCAGGTTCAGCTGCAGCAGTCTGGACCTGAGCTGGTGAAGCTG 111
QY 428 GGGCTCTAGTGAAGATTTCTGCTAGCAAGATTTCTGCTAGCAAGTCTGCTAGCAATTCAGTAGCTCTCTGGATGA 487
Db 112 GGGCTCTAGTGAAGATTTCTGCTAGCAAGTCTGCTAGCAATTCAGTAGCTCTCTGGATGA 171
QY 488 ACTGGGTGAAGCAGAGGCTCGACAGGCTCTTGAGTGAATTTGAGCGGATTTATCTGGAG 547
Db 172 ACTGGGTGAAGCAGAGGCTCGACAGGCTCTTGAGTGAATTTGAGCGGATTTATCTGGAG 230
QY 548 ATGGAGATTTCTAATCTACATCGGAATTTTCAGGCGCAAGCCATCTACTGACACAGCAAAAT 607
Db 231 ATGGAGATTTCTAATCTAATCGGAATTTTCAGGCGCAAGCCATCTACTGACACAGCAAAAT 290
QY 608 COTCCAGCAGACGCTTACATGCACTCAGCAGCTTACGCTTGTGGACTCTGTGGACTCTCGGCTATT 667
Db 291 COTCCAGCAGACGCTTACATGCACTCAGCAGCTTACGCTTGTGGACTCTGTGGACTCTCGGCTATT 350
QY 668 TCTGTGCAAGATCGGGTGTCTAGCTTATGCTATGCTAGTCTGGTCAAGGACCTCAG 727
Db 351 TCTGTGCAAGATCGGGTGTCTAGCTTATGCTATGCTAGTCTGGTCAAGGACCTCAG 407
QY 728 TCACCGTCTCCTC 740
Db 408 TCACAGTCTCCTC 420

```

```

RESULT 15
CA577519      417 bp mRNA linear EST 19-NOV-2002
LOCUS        K0710F01-SN NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1-)
DEFINITION   cDNA Library (Long) Mus musculus cDNA clone NIA:K0710F01
IMAGE:30074556 5', mRNA sequence.
ACCESSION    CA577519
VERSION      CA577519.1 GI:25125910
KEYWORDS     EST.
SOURCE       Mus musculus (house mouse)
ORGANISM     Mus musculus
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE       1 (bases 1 to 417)
JOURNAL     Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Luo, A., Carter, M.G.,
            Aiba, K., Taub, D., Longo, D.L., Keller, J. and Ko, M.S.H.
            Systematic Analyses of NIA Mouse Hematopoietic Stem Cell
            (Lin-/c-Kit-/Sca-1-) cDNA Library (Long)
            Unpublished (2001)
            Other ESTs: K0710F01-3
            Contact: Dawood B. Dudekula
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@igsun.grc.nia.nih.gov
            Plate: K0710 row: F column: 01
            Seq primer: M13 Reverse

```

High quality sequence stop: 417

POLYA=No.

## FEATURES

source

Location/Qualifiers

```

1. .417
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6NCr"
/db_xref="niaEST:K0710F01-5N"
/db_xref="taxon:10090"
/clone="NIA:K0710F01 IMAGE:30074556"
/tissue_type="Hematopoietic Stem Cell"
(Lin-/C-Kit-/Sca-1-)
/dev_stage="Age approx.10 weeks old"
/lab_hosts="DH10B"
/clone_lib="NIA Mouse Hematopoietic Stem Cell"
(Lin-/C-Kit-/Sca-1-) cDNA Library (Long)
/note="vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2:
NotI; Mouse cDNA project by the Laboratory of Genetics,
National Institute on Aging (NIA), Intramural Research
Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is
a long-transcript enriched cDNA library (Ref. Genome Res.
11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
obtained from Dr. Dennis Taub, Dan Longo (National Cancer
Institute on Aging, USA), Jonathan Keller (National Cancer
Institute, USA). Double-stranded cDNAs were synthesized
with an Oligo(dT) primer [Invitrogen:
5'-pGACTAGTCTAGATCGAGCGCGCCCTTTT-3'] from
0.9 ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to Lox-linker Lu-Sal4, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer Sal4-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pSPORT1 plasmid vector.
The DH10B E. coli host was transformed with the ligation
mixture by the standard chemical method. The average
insert size is about 2.1 kb. The library was constructed
by Yulan Piao (NIA)."
```

## ORIGIN

```

Query Match          36.1%; Score 267.8; DB 14; Length 417;
Best Local Similarity 91.3%; Pred. No. 1.1e-52;
Matches 284; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 368 CCGGTGGTGGTGGATCCAGAGTCCAACTGCAACAGTCTGGACCTGAGCTGGTGAAGCTG 427
Db      |||
QY 428 GGGCCTCAGTGAAGATTTCTTGCRAAGATTCCTGGCTAGCGATTCAATAGCTCTTGGATGA 487
Db      |||
QY 136 GGGCCTCAGTGAAGATTTCTTGCRAAGGCTTCTGGCTAGCGATTCAATAGCTCTTGGATGA 195
Db      |||
QY 488 ACTGGGTGAAGCAGAGGCTGACAGGGCTTTGAGTGGATTGGACGGATTTATCCTGGAG 547
Db      |||
QY 196 ACTGGGTGAAGCAGAGGCTGGAAGGGCTTTGAGTGGATTGGACGGATTTATCCTGGAG 255
Db      |||
QY 548 ATGGAGATTTCTTAATGGAATTCGAGGCGAAGCCATCTGACGACAGCAAAAT 607
Db      |||
QY 256 ATGGAGATACTAATGGAATTCGAGGCGAAGCCATCTGACGACAGCAAAAT 315
Db      |||
QY 608 CCTCCAGACACGCTTACATGAGCTCAGCAGGCTGACCTCTGTGACTCTCGGCTCTATT 667
Db      |||
QY 316 CCTCCAGACACGCTTACATGAGCTCAGCAGGCTGACATCTGAGGACTCTCGGCTCTACT 375
Db      |||
QY 668 TCTGTGCAAGA 678
Db      |||
QY 376 TCTGTGCAAGA 386
```

Search completed: April 26, 2004, 17:37:51

Job time : 2550 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2004, 15:43:02 ; Search time 393 Seconds  
(without alignments)

8009.961 Million cell updates/sec

Title: US-10-620-049-24

Perfect score: 741  
Sequence: 1 gatattcagatgacacagac.....cctcagtcacgctctctcctcg 741

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 337863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 674726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 29Jan04:\*

1: geneseqn1980s:\*

2: geneseqn1990s:\*

3: geneseqn2000s:\*

4: geneseqn2001as:\*

5: geneseqn2001bs:\*

6: geneseqn2002s:\*

7: geneseqn2003as:\*

8: geneseqn2003bs:\*

9: geneseqn2003cs:\*

10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	499.4	67.4	1803	3 AAAS8773	Aaas8773 DNA encod
2	499.4	67.4	1943	6 ABK13467	Abk13467 DNA encod
3	499.4	67.4	1944	7 ABX14413	Abx14413 DNA encod
4	482.8	65.2	2952	6 ABK13464	Abk13464 DNA encod
5	479.2	64.7	913	2 AAQ81076	Aaq81076 Bispecifi
6	463.8	62.6	1527	9 ADD25786	Add25786 Binding d
7	463.8	62.6	1687	9 ADE86042	Ade86042 DNA encod
8	431.8	58.3	2691	6 ABK13612	Abk13612 DNA encod
9	421	56.8	1953	8 ACF05482	Acf05482 Anti-CD7
10	421	56.8	1956	8 ACF05481	Acf05481 Anti-CD7
11	418.6	56.5	1457	3 AA39167	Aa39167 Apoptobod
12	406	54.8	1545	6 AAD32522	Aad32522 Murine CC
13	404.2	54.5	1698	3 AAZ43431	Aaz43431 Fv-antibo
14	395.8	53.4	1817	7 ACC79606	Acc79606 Plasmid p
15	394.2	53.2	1611	3 AAZ30332	Aaz30332 Nucleotid
16	391.6	52.8	795	6 ABK89855	Abk89855 Mouse sin
17	391.6	52.8	795	9 ADD25448	Add25448 Binding d
18	385.6	52.0	1637	3 AAZ89358	Aaz89358 Bispecifi
19	384.8	51.9	1065	2 AAT94963	Aat94963 R. pipien
20	384.8	51.9	1065	2 AAT94967	Aat94967 R. pipien
21	384.8	51.9	1074	2 AAT94965	Aat94965 R. pipien
22	384.8	51.9	1074	2 AAT94968	Aat94968 R. pipien
23	384.8	51.9	1086	2 AAT94966	Aat94966 R. pipien

24	384.8	51.9	1137	2 AAT94964	Aat94964 R. pipien
25	383.8	51.8	825	6 ABK89854	Abk89854 Mouse sin
26	383.8	51.8	825	9 AAD25447	Aad25447 Binding d
27	381.8	51.5	819	2 AAQ20381	Aaq20381 Sequence
28	379.2	51.2	752	4 AAF81912	Aaf81912 Single ch
29	379.2	51.2	782	4 AAF81913	Aaf81913 Single ch
30	373.2	50.4	1817	7 ACC79607	Acc79607 Plasmid p
31	371.2	50.1	1612	4 AAC86562	Aac86562 DNA encod
32	371.2	50.1	1614	8 AAT43497	Aat43497 huNR-LU-1
33	368	49.7	786	8 AAD56816	Aad56816 MT413 VL-
34	368	49.7	993	8 AAD56805	Aad56805 Chemokine
35	368	49.7	1017	8 AAD56800	Aad56800 Chemokine
36	367.8	49.6	848	2 AAQ41069	Aaq41069 26-10 sFv
37	363.8	49.1	2871	3 AAZ50586	Aaz50586 DC8scFv-e
38	361.8	48.8	1653	3 AAZ43432	Aaz43432 Fv-antibo
39	358.6	48.4	925	3 AAZ44205	Aaz44205 Murine de
40	358.6	48.4	925	3 AAZ58663	Aaz58663 Antibody
41	356.4	48.1	1299	2 AAV63620	Aav63620 Nucleic a
42	356.4	48.1	1299	2 AAZ09862	Aaz09862 pET-11d-E
43	356.4	48.1	1320	2 AAV63617	Aav63617 Nucleic a
44	356.4	48.1	1320	2 AAZ09859	Aaz09859 pET-11d-A
45	356.4	48.1	6727	2 AAV63616	Aav63616 Nucleic a

#### ALIGNMENTS

#### RESULT 1

AAAS8773  
ID AAAS8773 standard; DNA; 1803 BP.

XX AC AAAS8773;

XX DT 20-OCT-2000 (first entry)

XX DE DNA encoding an immunotoxin comprising a CD3 binding domain and PE28.

XX KW Immunotoxin; CD3-binding domain; Pseudomonas exotoxin A; immune system;

XX KW scFv(UCHT-1)-PE28; T-cell mediated disease; transplant rejection;

XX KW host versus graft disease; graft versus host disease;

XX KW bone marrow transplant; ss.

XX OS Synthetic.

OS Mus sp.

OS Pseudomonas aeruginosa.

XX PN WO200041474-A2.

XX PD 20-JUL-2000.

XX PF 13-JAN-2000; 2000WO-EP000245.

XX PR 15-JAN-1999; 99US-00232445.

PR 25-JAN-1999; 99US-00236968.

PR 07-OCT-1999; 99US-00414134.

XX PA (NOVS ) NOVARTIS AG.

PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.

XX PI Digan ME, Lake P, Wright RM;

XX DR WPI; 2000-482739/42.

DR P-PSDB; AAB07461.

XX PT Recombinant immunotoxin used for prophylaxis and treatment of T-cell

PT mediated diseases e.g. transplantation rejection.

XX PS Disclosure; Page 55-59; 75pp; English.

XX CC The present sequence encodes a recombinant immunotoxin comprising a CD3-

CC binding domain and a Pseudomonas exotoxin A component. It is designated

CC scFv(UCHT-1)-PE28. The immunotoxins of the invention are used for the

CC prophylaxis or treatment of T-cell mediated diseases or conditions of the

CC immune system. They are also used to condition a patient to be  
 CC transplanted with cells, or a tissue or an organ of a donor and for the  
 CC prophylaxis and/or treatment of acute or chronic transplantation  
 CC rejection, host versus graft disease and/or graft versus host disease in  
 CC a patient to undergo a bone marrow transplant, where the CD3-bearing cell  
 CC population in the patient is depleted and an inoculum comprising isolated  
 CC bone marrow and/or stem-cell enriched peripheral blood cells of the donor  
 CC treated with immunotoxin is inoculated into the patient  
 XX  
 SQ Sequence 1803 BP; 363 A; 575 C; 567 G; 298 T; 0 U; 0 Other;

Query Match 67.48; Score 499.4; DB 3; Length 1803;  
 Best Local Similarity 81.44; Pred. No. 1e-120;  
 Matches 610; Conservative 0; Mismatches 121; Indels 18; Gaps 2;

QY 1 GATATTGAGTACACAGAGTACATCTCCCTGTCTGCTCTCTGGGAGACAGAGTCAAC 60  
 DB 7 GACATCCAGATGACCCAGAGACACCTCTCTGCTGCTCTCTGGGAGACAGAGTCAAC 66  
 QY 61 GTGAGTTGCGAGGCAAGTACAGGACATAGGATTAATTTAACTGGTATCAGAGAACCA 120  
 DB 67 ATCAGTTGCGAGGCAAGTACAGGACATAGGATTAATTTAACTGGTATCAGAGAACCA 126  
 QY 121 GACGGAATCTGTTAAATTCCTGATCTACTACATCATCAAGATTATGCCAGGAGTCCCATCA 180  
 DB 127 GATGGAATCTGTTAAATTCCTGATCTACTACATCATCAAGATTATGCCAGGAGTCCCATCA 186  
 QY 181 AGTTTCAGTGCAGTGGTCTGGAACAGATTAATTCCTCACCATTAAACCTGGAGCAG 240  
 DB 187 AAGTTTCAGTGCAGTGGTCTGGAACAGATTAATTCCTCACCATTAAACCTGGAGCAG 246  
 QY 241 GAAATATTGGCACTTACTTTTGGCAACAGGCAATACGCTCCGTCGAGCGTTCCGGTGA 300  
 DB 247 GAGATATTGCCATCTACTTTTGGCAACAGGCAATACGCTCCGTCGAGCGTTCCGGTGA 306  
 QY 301 GGCACCAAGCTGGAAATAAAGCTGGTGGAGTGGTCTGATGATGGTGGTGGTGGTGG 360  
 DB 307 GGCACCAAGCTGGAAATAAAGCTGGTGGAGTGGTCTGATGATGGTGGTGGTGGTGG 366  
 QY 361 GCGCGCTCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 420  
 DB 367 AGC-----GGTGGCGGATCTGAGTGGAGTCCAGCTCCAGCTCCAGCTCCAGCTGG 417  
 QY 421 AAGCTGGGGCTCAGTGAAGATTTCTGCAAGATTTCTGCTACGATTCAGTCAATCAATGCT 480  
 DB 418 AAGCTGGAGCTTCAATGAAGATATCTGCAAGGCTTCTGCTACTCACTCACTGGCTAC 477  
 QY 481 TGGATGACTGGTGCAGCAGGCTGGAGCGGTCTGAGTGGATTTGAGCGGATTTAT 540  
 DB 478 ACCATGAATGGTGAAGCAGGATCTGGAAGAACCTTGGTGGATGGGACTTATTAAT 537  
 QY 541 CCTGGAGATGGAGATTTCTAATAAGTGGAAATTCGAGGCGCAAGGCCATCTGACAGCA 600  
 DB 538 CCTTACAAAGGTGTGTAGTACCTACACCAAGGTTCAAGGACAAAGGCCACATTAAGTGA 597  
 QY 601 GACAAATCTCCAGCAGAGCTACATGAGCTCAGCAGCTGAGCTCTGTGGACTCTGG 660  
 DB 598 GACAAATCTCCAGCAGAGCTACATGAGCTCAGCAGCTGAGCTCTGTGGACTCTGG 657  
 QY 661 GTCTATTTCTGCAAGATCGGGGTGTGCTAGCTTATGCT-----ATGCACTATGG 711  
 DB 658 GTCTATTTCTGCAAGATCGGGGTGTGCTAGCTTATGCT-----ATGCACTATGG 717  
 QY 712 GGTCAAGAACCTCAGTCAAGCTCTCTCTC 740  
 DB 718 GCGCAGGAGCAACCGTCCCGTCTCTCTC 746

RESULT 2  
 ABK13467  
 ID ABK13467 standard; DNA; 1943 BP.  
 XX  
 AC ABK13467;

XX 23-APR-2002 (first entry)  
 DT DNA encoding Immunotoxin fusion protein DT389.sFv(UCHT1).  
 DE  
 XX Immunotoxin; pseudomonas exotoxin A; ERA; diphtheria toxin; DT;  
 KW cytostatic; immunosuppressive; immunostimulant; antidiabetic;  
 KW antirheumatic; antithrombotic; anti-HIV; anti-inflammatory;  
 KW anti-T cell immunotoxin fusion protein; antibody; UCHT1;  
 KW (Gly4Ser)3 linker; T cell leukaemia; lymphoma; gene; ds;  
 KW graft-versus-host disease; Fv; autoimmune disease; transplant rejection;  
 KW systemic lupus erythematosus; type I diabetes; rheumatoid arthritis;  
 KW myasthenia gravis; multiple sclerosis; AIDS; DT389.sFv(UCHT1);  
 KW acquired immunodeficiency syndrome; chronic immunosuppression.  
 XX  
 OS Mus sp.  
 OS Corynebacterium diphtheriae.  
 OS Synthetic.  
 OS Chimeric.  
 XX  
 XX Key Location/Qualifiers  
 CDS 3..1932  
 /\*tag= a  
 /product= "DT389.sFv(UCHT1)"  
 FT WO200187982-A2.  
 FN  
 XX  
 PD 22-NOV-2001.  
 XX  
 PF 18-MAY-2001; 2001WO-US016125.  
 XX  
 PR 18-MAY-2000; 2000US-00573797.  
 XX  
 XX (NOVS ) NOVARTIS AG.  
 PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 XX Neville DM, Thompson JT, Hu H, Woo J, Ma S, Hexham JM, Digan ME;  
 XX WPI; 2002-121980/16.  
 DR P-PSDB; AAU75382.  
 XX  
 XX New anti-T cell immunotoxin fusion protein comprising a truncated  
 PT diphtheria toxin moiety, a connector, and one single chain Fv of the  
 PT variable region of a UCHT1 antibody, useful for treating e.g. autoimmune  
 PT diseases.  
 XX  
 PS Claim 34; Fig 35; 307pp; English.  
 XX  
 CC The invention relates to an anti-T cell immunotoxin fusion protein,  
 CC comprising from the amino terminus, a truncated diphtheria toxin moiety,  
 CC a connector, and one single chain Fv of the variable region of a UCHT1  
 CC antibody. The single chain Fv comprises VL, L, or VH, L is a Gly-Ser  
 CC linker, and VL and VH are the variable light and heavy domains of the  
 CC anti-CD3 antibody UCHT1. Also included are a nucleic acid encoding the  
 CC novel immunotoxin fusion protein, a vector comprising the nucleic acid  
 CC and a cell comprising the nucleic acid. The immunotoxin may also be a  
 CC pseudomonas exotoxin A (ERA). The immunotoxins are useful for treating T  
 CC cell leukaemias or lymphomas, graft-versus-host diseases, and autoimmune  
 CC diseases by inducing immune tolerance. The immunotoxin fusion proteins  
 CC may be used in vivo to systemically reduce populations of T cells, or ex  
 CC vivo to effect T-cell depletion from a treated cell population. The  
 CC fusion proteins can be administered to a subject who is or will be a  
 CC recipient of an allograft to prevent or reduce T-cell mediated acute  
 CC or chronic transplant rejection of the transplanted allogeneic cells,  
 CC tissue or organ in the subject as well as treat other T-cell mediated  
 CC diseases such as systemic lupus erythematosus, type I diabetes,  
 CC rheumatoid arthritis, myasthenia gravis, multiple sclerosis, infectious  
 CC diseases of the immune system (e.g. AIDS (acquired immunodeficiency  
 CC syndrome)) and chronic immunosuppression. The present sequence encodes an  
 CC immunotoxin fusion protein of the invention comprising 389 residues from  
 CC the N-terminal glycine of mature DT toxin, a (Gly4Ser)3 linker and one  
 CC single Fv chain from antibody UCHT1, DT389.sFv(UCHT1)







```

Db 1675 GAGGATATGCCACTTACTTTTCCCAACAGGGTAATACGCTTCGGTGGAGCTTCGCTGGA 1734
Qy 301 GGCACCAAGCTGGAATAAACAAGCTGGTGGAGTGGTCTGTGATGGTGGTCTGTGGCGGC 360
Db 1735 GGCACCAAGCTGGAATAAAGAGGCGGAGGCAG-----CGGAGGC 1776
Qy 361 GCGGGCTCGGTGGTGGTGGATCCGAGGTCCTCAACTGCAAGCTGCAAGCTGAGCTGAGCTGGTG 420
Db 1777 GGTGGCTCGGAGGGGAGGCTCGGAGGTGCAGCTCCAGCAGCTCTGCACTGAGCTGGTG 1836
Qy 421 AAGCCTGGGGCCTCAGTGAAGATTTCTGCAAAAGATTTCTGGTACCGCATTCATAGCTCT 480
Db 1837 AAGCCTGGAGCTTCAATGAAGATATCTGCAAGGCTTCTGGTTACTCAATTCATCTGCTAC 1896
Qy 481 TGGATGAACCTGGTGAAGCAGAGGCTCGACAGGGCTTGGTGGATTCGACGGATTTAT 540
Db 1897 ACCATGAACCTGGTGAAGCAGAGATCTGGAAGAACCTTGGTGGATGGGACTTATAT 1956
Qy 541 CCTGGAGATGGAGATTTAACTAACTGGGAATTCGAGGGCAAGGCCATACGACAGCA 600
Db 1957 CCTTACAAAGGTGTAGTACTCAACCAAGAGTTCAGGACAAGGCCACATTAACCTGA 2016
Qy 601 GACAAATCCTCCAGCAGAGCTACATGACGCTCAGAGCTGACCTCTGGACTCTGG 660
Db 2017 GACAAGTCACTCCAGCAGAGCTACATGGAACCTCTCAGCTGACATCTGAGGACTCTGCA 2076
Qy 661 GTCTATTCTGTGCAAGATCGGGGTGCTACTAGTTATGCTATGACTACTGGGCTCAAGGA 720
Db 2077 GTCTATTACTGTCAAGATCGGGGTACTACGGTGTAGTACTGCTGATGCTGG 2136
Qy 721 ACTCAGTCAAC 732
Db 2137 GCGCAGGCAAC 2148

RESULT 5
AAQ81076
XX AC AAQ81076 standard; cDNA; 913 BP.
XX AC AAQ81076;
XX AC AAQ81076;
DT 25-MAR-2003 (revised)
DT 14-MAR-1995 (first entry)
XX AC AAQ81076;
DE Bispecific CD3-L6FvIg antibody derivative coding sequence.
XX fusion protein; recombinant bispecific single chain antibody;
XX helical peptide linker; anti-L6 antibody; tumour cell antigen;
XX anti-CD3 antibody; variable region; ss.
XX OS Synthetic.
XX Key Location/Qualifiers
FH 7..912
FT CDS /*tag= b
FT sig_peptide /*tag= a
FT misc_RNA /*tag= c
FT /*tag= "L6 light chain variable region leader"
FT /*tag= "encodes CD3 VL-VH fusion"
FT /*tag= "encodes CD3 VL-VH fusion"
FT /*tag= "encodes hinge region of Fc domain"
FT /*tag= "encodes helical peptide linker"
XX EP610046-A2.
XX PD 10-AUG-1994.
XX PF 31-JAN-1994; 94EP-00300692.

```

```

XX 01-FEB-1993; 93US-00013420.
PR 13-SEP-1993; 93US-00121054.
XX
PA (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI Ledbetter JA, Gilliland LK, Hayden MS, Linsley PS, Bajorath J;
PI Fell PH;
XX WPI; 1994-250885/31.
DR P-PSDB; AAR60206.
XX
PT Expression vector encoding bispecific fusion protein - having binding
PT domains for separate targets joined by helical peptide, useful e.g. for
PT diagnosis and treatment.
XX
PS Example 1; Fig 11; 50pp; English.
XX
CC The VL and VH sequences of the anti-CD3 hybridoma G19-4 were amplified by
CC PCR methods. A gene fusion was constructed from the two amplified domains
CC and a (Gly4Ser)3 linker. The amino terminus of the VL-VH fusion cassette
CC was fused at the SalI site to the L6 light chain variable region leader
CC peptide and the carboxy-terminus was fused directly to the hinge region
CC of the Fc domain at the BclI site and/or to a short "helical" peptide
CC linker to construct the bispecific CD3-L6FvIg antibody derivative. The
CC variable regions for L6 were fused in frame to the opposite end of the
CC helical linker (not included in AAQ81076). (Updated on 25-MAR-2003 to
CC correct FN field.)
XX
SQ Sequence 913 BP; 252 A; 223 C; 228 G; 210 T; 0 U; 0 Other;
Query Match 64.7%; Score 479.2; DB 2; Length 913;
Best Local Similarity 79.4%; Pred. No. 1.7e-115;
Matches 581; Conservative 0; Mismatches 148; Indels 3; Gaps 1;
Qy 1 GATATTTCAGATGACACAGACTACATCTCCCTCTGCTGCTCTCTCTGGGAGACAGAGTCAAC 60
Db 76 GACATCCAGATGACACAGACTACATCTCCCTCTGCTGCTCTCTCTGGGAGACAGAGTCAAC 135
Qy 61 GTCAGTTGAGGGCAAGTCAGGACATTTAGGAATTTATTTAACTGGTATCAGCAGAAACCA 120
Db 136 ATCAGTTGAGGGCAAGTCAGGACATTTGCAATTTATTTAACTGGTATCAGCAGAAACCA 195
Qy 121 GACGGAATCTGTTAAATCTCTGATCTACTACATCAAGATTACTGCCAGGAGTCCCATCA 180
Db 196 GATGGAATCTGTTAAATCTCTGATCTACTACATCAAGATTACTCGAGGATCCCATCA 255
Qy 181 AGGTTTCAGTGGCAGTGGTCTGGAAACAGATTATTTCCCTCACCATTTAACTGGAGCAG 240
Db 256 AGGTTTCAGTGGCAGTGGTCTGGAAACAGATTATTTCTCAGCATTTGCCAACCTGCAACCA 315
Qy 241 GAAGATATTTGGCAGTACTTTTGGCAACAGGCAATAGCCTCGGTGAGCTTCGGTGGGA 300
Db 316 GAAGATATTTGGCAGTACTTTTGGCAACAGGCAATAGCCTCGGTGAGCTTCGGTGGGA 375
Qy 301 GGCACCAAGCTGGAATAAAGCTGGTGGAGTGGTCTCTGATGGTGGTGGTCTCGCGCGC 360
Db 376 GGCACCAAGCTGGTAAACCAAGCGGAGCTCGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 435
Qy 361 GCGGCTCGGTGGTGGTGGATCCGAGGTCCAACTGGCAACAGCTCTGGACCTGAGCTGGT 420
Db 436 GGTGG---CGGCGGATCTATCGATGAGTCCAGCTGCAACAGCTCTGGACCTGAACTGGTG 492
Qy 421 AAGCCTGGGGCCTCAGTGAAGATTTCTGCAAAAGATTTCTGGCTAGCATTCATAGCTCT 480
Db 493 AAGCCTGGAGCTTCAATGACATGTCTTGCAGGCGCTCTGGTTACTCATTCATCGCTAC 552
Qy 481 TGGATGAATGGTGAAGCAGAGGCTGGACAGGCTCTTGGTGGATTTGGACGGATTAT 540
Db 553 ATCGTGAATGGTGAAGCAGAGGCTGGACAGGCTCTTGGTGGATTTGGACGGATTAT 612
Qy 541 CTGGGAGATGGAGATTCTAACTACAAATGGGAATTCGAGGGCAAGGCCATCTGACAGCA 600

```



XX 26-MAR-2002; 2002US-00107991.  
 XX (GENE-) GENE-CRAFT INC.  
 XX Ledbetter JA, Hellstrom I, Hayden-Ledbetter M, Hellstrom KE;  
 XX WPI; 2003-804045/75.  
 XX DR P-PSDB; ADB86039.  
 XX New culture system for generating tumor-reactive T-lymphocytes comprises  
 PT T-cells from a cancer patient, antigen-presenting cells, autologous or  
 PT allogeneic tumor cells, and immobilized antibodies to T-cell receptors.  
 XX Example 2; SEQ ID NO 1; 112pp; English.  
 XX The present sequence is of a DNA construct encoding a mouse-human  
 CC chimeric single chain (sc) Fv comprising the G19-4 mouse anti-human CD3  
 CC VL domain joined via a peptide linker to the G19-4 mouse anti-human CD3  
 CC VH domain, human IgG1 FC domain (hinge, CH2, CH3) and human CD80  
 CC transmembrane and cytoplasmic domains, with an N-terminal L6 V kappa  
 CC leader peptide. The coding sequence was constructed to provide expression  
 CC of active anti-CD3 monoclonal antibody scFv at the tumour cell surface.  
 CC The DNA construct, or a cell transfected with it, can be used for in vivo  
 CC cancer therapy. It is believed that the anti-CD3 scFv expression on the  
 CC surface of tumour cells induces polyclonal T cell activation and tumour  
 CC cell destruction, releasing tumour antigens and promoting a transition to  
 CC antigen-specific tumour immunity, detected as rejection of 'wild-type'  
 CC (not transfected) cells from the same tumour. In an example from the  
 CC invention, cells expressing anti-CD3 scFv at their surface were shown to  
 CC induce polyclonal stimulation of T cells to proliferate, produce Th1 type  
 CC lymphokines, to become cytolytic and to have antitumour activity in vivo.  
 XX SQ Sequence 1687 BP; 434 A; 470 C; 431 G; 352 T; 0 U; 0 Other;  
 Query Match 62.6%; Score 463.8; DB 9; Length 1687;  
 Best Local Similarity 79.3%; Pred. No. 2.2e-111;  
 Matches 594; Conservative 0; Mismatches 137; Indels 18; Gaps 3;  
 1 GATATTGAGTACACAGATCATATCTCCCTGCTCTGCTCTGCGGACAGAGTCACC 60  
 76 GACATCCAGATGACACAGATCATATCTCCCTGCTCTGCTCTGCGGACAGAGTCACC 135  
 61 GTGAGTTCAGGCGCAAGTCAGGACATAGGAAATTTAACTGGTATCAGCAGAAACCA 120  
 136 ATGAGTTCAGGCGCAAGTCAGGACATAGGAAATTTAACTGGTATCAGCAGAAACCA 195  
 121 GACGGAATCTTTAAATCTCTGATCTACTACATCAAGATTAATCTGCGGAGTCCCATCA 180  
 196 GATGGAACTGTTAAATCTCTGATCTACTACATCAAGATTAATCTGCGGAGTCCCATCA 255  
 181 AGGTCAGTGGCAGTGGTCTGAAACAGATTAATCTGCTCACCATTAAACCTGGAGCAG 240  
 256 AGGTCAGTGGCAGTGGTCTGAAACAGATTAATCTGCTCACCATTGCGGAGTCCCATCA 315  
 241 GAAGATATTGCGACTTACTTTTCCCAACAGGCAATACGCTCCGTTGGAGCTTCGGTGA 300  
 316 GAAGATATTGCGACTTACTTTTCCCAACAGGCAATACGCTCCGTTGGAGCTTCGGTGA 375  
 301 GGCACCAAGCTGGAATAAAACGTTGGAGTGGTCTGATGGTGGTCTGCTGGCGGC 360  
 376 GGCACCAAACTGGTAACCAAAACGAGTCCGTTGGCGGTGGCTCGGCGGTGGTGGTGG 435  
 361 GCGCGCTCCGTTGGTGGTGGTCCGAGTCCCACTGCAAGCTGCACTGAGCTGGTGG 420  
 436 GGTGG---CGCGGATCTATCGATAGGTCCAGCTGCAACAGCTGCACTGAGCTGGTGG 492  
 421 AAGCCTGGGCGCTCAGTGAAGATTTCTGCAAGATTTCTGGTACGCATTCATAGCTCT 480  
 493 AAGCCTGGAGCTTCAATG-----TCTGCAAGCGCTCTGGTTACTTCACTTGGCTAC 546  
 481 TGAATGACTGGTGAAGCAGAGCGCTGGACAGGCTTTCAGTGGATTTGACGAGTTAT 540

Db 547 ATCGTGAAGTGGCTGAAGCAGAGCCATGGAAAGAACTTGAAGTGGATTGAGCTTATTAAT 606  
 Qy 541 CTTGGAGATGGAGATTCTTAATCAATGGGAAATTCGAGGCGCAAGCCATATGACAGCA 600  
 Db 607 CCATACAAAGGCTCTTACTACTACCAACGAAATTCAGAGGCAAGGCCACATTAATGTA 666  
 Qy 601 GACAAATCTCCAGCAGACAGCTACATGACCTCAGAGCTCAGACCTCTGTGGACTCTGCG 660  
 Db 667 GACAAGTCATCCAGCAGACAGCTACATGAGCTCTCTAGTCTGACATCTGAAGACTCTGCA 726  
 Qy 661 GTCTATTCTTGTGCAAGATCGGGTCTCTAGCTTATGC-----TATGGACTACTGG 711  
 Db 727 GTCTATTACTGTGCAAGATCTGGGTACTATGGTGTACTCGGACTGGTACTTCGATGTCGG 786  
 Qy 712 GGTCAAGAACCTCAGTCACCGTCTCCTC 740  
 Db 787 GGCACAGGACCCAGGTCACCGTCTCCTC 815  
 RESULT 8  
 ABK13612  
 ID ABK13612 standard; DNA; 2691 BP.  
 XX AC ABK13612;  
 XX DT 23-APR-2002 (first entry)  
 XX DE DNA encoding an Immunotoxin fusion protein.  
 XX Immunotoxin; pseudomonas exotoxin A; ETA; diphtheria toxin; DT;  
 KW cytosolic; immunosuppressive; immunostimulant; antidiabetic;  
 KW antirheumatic; antiarthritic; anti-HIV; anti-inflammatory;  
 KW anti-T cell immunotoxin fusion protein; antibody; UCHL1;  
 KW (Gly4Ser)3 linker; T cell leukaemia; lymphoma; gene; ds;  
 KW graft-versus-host disease; Fv; autoimmune disease; transplant rejection;  
 KW systemic lupus erythematosus; type I diabetes; rheumatoid arthritis;  
 KW myasthenia gravis; multiple sclerosis; AIDS;  
 KW acquired immunodeficiency syndrome; chronic immunosuppression.  
 XX Mus sp.  
 OS Corynebacterium diphtheriae.  
 OS Synthetic.  
 OS Chimeric.  
 XX WO2000187982-A2.  
 XX 22-NOV-2001.  
 XX 18-MAY-2001; 2001WO-US016125.  
 XX 18-MAY-2000; 2000US-00573797.  
 XX (NOVS ) NOVARTIS AG.  
 PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX Neville DM, Thompson JT, Hu H, Woo J, Ma S, Hexham JM, Digan ME;  
 DR WPI; 2002-121980/16.  
 XX New anti-T cell immunotoxin fusion protein comprising a truncated  
 PT diphtheria toxin moiety, a connector, and one single chain Fv of the  
 PT variable region of a UCHL1 antibody, useful for treating e.g. autoimmune  
 PT diseases.  
 XX Claim 91; Page 300-301; 307pp; English.  
 CC The invention relates to an anti-T cell immunotoxin fusion protein,  
 CC comprising from the amino terminus, a truncated diphtheria toxin moiety,  
 CC a connector, and one single chain Fv of the variable region of a UCHL1  
 CC antibody. The single chain Fv comprises VL, L, or VH, L is a Gly-Ser  
 CC linker, and VL and VH are the variable light and heavy domains of the  
 CC anti-CD3 antibody UCHL1. Also included are a nucleic acid encoding the

novel immunotoxin fusion protein, a vector comprising the nucleic acid and a cell comprising the nucleic acid. The immunotoxin may also be a pseudomonas endotoxin A (ETA). The immunotoxins are useful for treating T cell leukaemias or lymphomas, graft-versus-host diseases, and autoimmune diseases by inducing immune tolerance. The immunotoxin fusion proteins may be used in vivo to systemically reduce populations of T cells, or ex vivo to effect T-cell depletion from a treated cell population. The fusion proteins can be administered to a subject who is or will be a recipient of an allotransplant to prevent or reduce T-cell mediated acute or chronic transplant rejection of the transplanted allogeneic cells, tissue or organ in the subject as well as treat other T-cell mediated diseases such as systemic lupus erythematosus, type I diabetes, rheumatoid arthritis, myasthenia gravis, multiple sclerosis, infectious diseases of the immune system (e.g. AIDS (acquired immunodeficiency syndrome)) and chronic immunosuppression. The present sequence encodes an immunotoxin fusion protein of the invention comprising 390 residues from the N-terminal glycine of mature DT toxin, a (Gly4Ser)3 linker and one single Fv chain from antibody UCHL1, and has had its AT content altered to enable expression in yeast

Sequence 2691 BP: 669 A; 649 C; 701 G; 672 T; 0 U; 0 Other;

Query Match 58.3%; Score 431.8; DB 6; Length 2691;  
Best Local Similarity 76.1%; Pred. No. 6.1e-103;  
Matches 570; Conservative 0; Mismatches 152; Indels 27;

Qy	1	GATATT	CAGATG	CACAG	CTACAT	CTCTCC	TGTC	TGCTCC	TCTCTG	GAGACAG	AGTCA	CACC	60
Db	1180	GACATC	CAGATG	ACCCAG	ACCACT	CTCTCC	TGTC	TGCTCC	TCTCTG	GAGACAG	AGTCA	CACC	1239
Qy	61	GTCA	TTGTCAGG	CGCAAGT	CAAGG	CAATAGG	AAATTA	TAAACT	TGGTAT	CAGCAG	AAACCA	120	
Db	1240	ATCA	TTGTCAGG	CGCAAGT	CAAGG	CAATAGG	AAATTA	TAAACT	TGGTAT	CAGCAG	AAACCA	1299	
Qy	121	GACGG	AATCTGT	TAAATTC	TGATCTA	CA	CATCAA	AGATTAC	TG	CCAGGAGT	TCCCATCA	180	
Db	1300	GACGGT	ACTGTCA	AGTTGTTG	ATCTACT	ACACTT	CCAGATTG	CACTCT	GGTGTGCC	ATCT	1359		
Qy	181	AGTT	CAGTGC	AGTGGGT	CTGGA	ACAGATTAT	TCTCCT	CACCA	TAAACACCT	TGGAGCAG	240		
Db	1360	AGTT	CTCTG	GAATCGTT	CTGTACT	GA	CTACTCCT	TGACCA	TCTCCA	ACTTGGAGCA	1419		
Qy	241	GAAG	ATATGG	CACTTACT	TTTGG	CAACAGG	GCNAATAC	GCCT	CGTGG	AGAGTTGCGTGGGA	300		
Db	1420	GAGG	ATATGG	CACTTACT	TTTGG	CAACAGG	GCNAATAC	GCCT	CGTGG	AGAGTTGCGTGGGA	1479		
Qy	301	GGC	ACCAAGCT	TGGA	ATAA	ACG	TGTTGG	TGAGGT	TGTTCT	GATGGTGGTGTCT	TGGCGGC	360	
Db	1480	GGC	ACCAAGTT	TGGAGATCA	AGGTGG	AGGA	-----	GGTTCT	GAGGT	1521			
Qy	361	GGCG	GCTCCGGT	TGTTGG	TATCCG	AGGTCCA	ACTG	CAACAGT	CTGGA	CCTGAGCTGGTG	420		
Db	1522	GGT	GAATCTG	GAGTGG	AGGTCT	GAGTTCA	ATTG	CAACAA	CTGGA	CCTGAGCTGGTG	1581		
Qy	421	AAG	CTGGG	GCTCAGT	GAAGATT	TCTG	CA	AAATTC	TGGTAC	GATCAATAGCTCT	480		
Db	1582	AAG	CTGGAG	CTTCA	ATTGA	GATATCT	CTG	CAAGCTTCT	GGTACT	CATTCACTGGGTAC	1641		
Qy	481	TGG	ATGA	ACTGGT	GAAGCAG	AGCCTGG	ACAG	GGTCTT	GAGT	TGGAATTGACGGAATTTAT	540		
Db	1642	ACC	ATGA	ACTGGT	GAAGCAG	AGTCA	TG	GA	AAACCTT	GAGTGGATGGGTTTGATTAA	1701		
Qy	541	CCT	GGAG	ATGAGA	TTCTTA	CTCA	CA	ATGGG	AAATTC	GAGGGCAGG	CCATACTGACAGCA	600	
Db	1702	CTT	TACA	AGGGTGT	CTG	AC	CTA	CA	ACAG	AAATTC	AAGGA	CAAGGCTACTTTCACTGTA	1761
Qy	601	GACA	AAATCCT	CCAG	CACAG	CC	TAC	ATG	CAGCT	CTCAG	CGCTG	CACTCTGGC	660
Db	1762	GAC	AGTCA	TCCAG	CACAG	CC	TAC	ATG	AAATCT	CTCAG	CTG	CACTCTGCA	1821
Qy	661	GTCT	ATTTCT	GTG	CAAGAT	CCGG	GTGCTA	CGTTAT	GC	-----	ATGCA	TACTGG	711
Db	1822	GTCT	ATTTCT	GTG	CAAGAT	CCGG	GTGCTA	CGGTG	ATGCT	CGGTACT	CGGTACT	CGGTCTGG	1881

QY 712 GGTCAGGAACCTCAGTACC GTCTCTC TC 740  
||| ||| ||| ||| ||| ||| ||| |||  
pb 1882 GGTGCTGGTACTCTGTCA CTGTCTCTC TC 1910

## RESULT 9

ACF05482 standard: cDNA: 1953 bp.

AC ACF05482;

DT 06-NOV-2003 (first entry)

XX Anti-CD7 antibody TH-69 scFv-Pseudomonas exotoxin A fusion sequence. DE

XX CD7; antibody; TH-69; scFv; exotoxin A; immunotoxin; cytostatic;  
KW immunosuppressive; vaccine; gene; ss.

XX  
OS  
Mus sp.

OS *Pseudomonas* sp.  
OS synthetic.

OS Synthetic.  
OS Chimeric.

XX	Key	Location/Qualifiers
FH		

```

ET CDS
T. .195.
/*tag=

```

113 partial /  
114 partial /  
115 partial /

```

ET /particular
ET /product= "TH-69 scFv-E

```

## References

PN WO2003051926-A2.

26-JUN-2003

11-DEC-2002; 2002WO-EP014064.

PR 14-DEC-2001; 2001US-0339422P.

PA (UYER-) UNIV ERLANGEN-NUERNBURG.

PI Fev GHM, Gramatzki M, Peipp M;

XX  
DR WPI: 2003-523519/49.

DR P-PSDB; ABR62591.

New fusion protein comprising functionally linked components of an anti-CD7 antibody or its fragment and an immunotoxin, useful for treating CD7 positive acute T-cell and/or myeloid leukemias, or Graft-versus-host disease.

PS Claim 11: Page 10-11: 55pp: English.

The present sequence encodes a novel fusion protein comprising an scFv fragment of anti-CD7 monoclonal antibody RH-69 and domains II and III of Pseudomonas exotoxin A (ETA), joined via a peptide linker, and including an N-terminal 6xHis tag and C-terminal KDEL sequence. This novel fusion protein binds to CD7-positive T-lymphoid cells and kills them by the induction of apoptosis. It is an example of anti-CD7 scFv immunotoxin fusion proteins of the invention. An expression vector, host cells, a method for producing the fusion protein, and a vaccine comprising the fusion protein are claimed. The fusion protein is used in the treatment of disorders involving a hyperproliferation of CD7-positive cells, especially acute T-cell and/or myeloid leukaemia, and also in the treatment or prophylaxis of Graft-versus-host disease. A claimed method of conditioning an animal (e.g. a human) to be transplanted with donor cells, tissue or organ comprises: (a) depleting the CD7-bearing cell population in the animal; (b) providing a transplant comprising isolated bone marrow and/or stem cell-enriched peripheral blood cells of the donor, where the CD7-bearing cell population in (a) and/or the transplant in (b) are treated with a T-cell depleting effective amount of the fusion protein or therapeutic composition comprising the fusion protein; and (c) introducing the transplant into the animal.

XX	SQ	Sequence	1953 BP	353 A	614 C	655 G	331 T	0 U	0 Other
		Query Match	56.8%	Score 421	DB 8	Length 1953			
		Best Local Similarity	74.9%	Pred. No. 3.7e-100					
		Matches 555	Conservative	0	Mismatches 180	Indels	6	Gaps	2
Qy	1	GATATTCAGATGACACAGACTACATCTCCTCTCTGCTCTCTCTGGAGACAGATCCACC	60						
Db	49	GATATCCAGATGACACAGACTACATCTCCTCTGCTGCTCTCTGGAGACAGATCCACC	108						
Qy	61	GTCACTTCGAGGGCAAGTCAGGACATTAGGAATATTAAACTGGTATCAGCAGAAACCA	120						
Db	109	ATCACTTGCAGTGCAGTCAGGCATTAGCAATATTATTAACTGGTATCAGCAGAAACCA	168						
Qy	121	GACGGAACTGTTAAATTCCTGATCTACTACACATCAAGATTACTGCCAGGATCCCATCA	180						
Db	169	GATGGAACCTGTTAAACTCCTGATCTATTACACATCAAGTTTACATCAGGATCCCATCA	228						
Qy	181	AGGTTTCAGTGGCAGTGGTCTTGGAACAGATTATTCCTCACCAATTAAACACTGGAGCAG	240						
Db	229	AGGTTTCAGTGGCAGTGGTCTTGGAACAGATTATTCCTCACCAATCAGCAACTCTGGAACCT	288						
Qy	241	GAAATATTGGCACTTACTTTTGGCAACAGGGCAATAGCCTCCGTGGAGCTTCGGTGA	300						
Db	289	GAAATATTGGCACTTATTTATTTGCACAGTATAGCAAGCTTCGTACAGTTCGGAGGG	348						
Qy	301	GGCACCAAGCTGGAATAAAAAAGTGGTGAGGTTGGTTCTGATGGTGGTGGTCTGGCGGC	360						
Db	349	GGGACCAAGCTGGAATAAAAAAGTGGTGAGGTTGGTTCTGATGGTGGTGGTCTGGCGGC	408						
Qy	361	GGCGGCTCCGGTGGTGGTGGATTCGAGAGTTCCACTGCACAGTCTGGAACCTGAGCTGGTG	420						
Db	409	GGCGGCTCCGGTGGTGGTGGATTCGAGAGTTCGCACTGGTGGAGTCTGGGGGAGGCTTAGTG	468						
Qy	421	AGACCTGGGGCCTCAGTGAAGATTTCCTGCAAGATTCTGCTACGATTCATAGCTCTCT	480						
Db	469	AGACCTGGGGGCTCCCTGAAACTCTCTGTGCACCTCTGGAATCACTTCAGTAGCTAT	528						
Qy	481	TGGATGAATCGGGTGAAGCAGAGGCTTGGAACAGGGTCTTGAGTGGATTGGAACGATTTAT	540						
Db	529	GCCATGCTCTGGGTTTCGCGACACTCCAGAGAAGAGGCTGGAGTGGGTCGATCCATTAGT	588						
Qy	541	CTTGAGATGGAGATTCTAACTACAATGGGAAATTCGAGGGCAAGGCCATCTATGCACGCA	600						
Db	589	AGTGG---TGGTTTCACTACTATCCAGACAGTGTGAAGGGCCGATTACCATCTCCAGA	645						
Qy	601	GACAAATCTCCAGCACACGCTTACATGACGCTCAGCAGCCTGACCTCTGTGAGCTCTGGG	660						
Db	646	GATAATGCCAGGAACATCTGTATCTGCAATGAGCAGTCTGAGGCTCTGAGGACACGGCC	705						
Qy	661	GTCTATTTCTGTGCAAGATCCGGGTTGCTACGTTTATGCTATGACTACTGGGGTCAAGGA	720						
Db	706	ATGTATTACTGTGCAAGAGACAGAGTACGGGGGTTA---CCTCGATGTCTGGGGCGCAGGG	762						
Qy	721	ACCTCAGTCAACGCTCTCTCTCG	741						
Db	763	ACCACGGTCAACGTTTCTCTCG	783						

## RESULT 10

RESOL1 TO  
ACF05481  
ID ACF05481 standard; cDNA; 1956 BP.

AC ACF05481;

06-NOV-2003 (first entry)

DE Anti-CD7 antibody TH-69 scFv-Pseudomonas exotoxin A fusion sequence.

XX  
KW CD7; antibody; TH-69; scFv; exotoxin A; immunotoxin; cytostatic;

immunosuppressive; vaccine; gene; ss

XX

W

OS	Mus sp.
OS	Pseudomonas sp.
OS	Synthetic.
OS	Chimeric.
XX	
Key	Location/Qualifiers
PH	1..1956
FT	/tag= a
FT	/partial
FT	/product= "TH-69 scFv-ETA fusion protein"
FT	/note= "No start codon"
FT	
CDS	

WO2003051926-A2.

26-JUN-2003.

11-DEC-2002: 2002WO-EP014064.

14-DEC-2001: 2001US-0339422P.

(UYER-) UNTV ERLANGEN-NÜRNBERG.

Prof. CHM Gramatzi M. Peinn M.

WPT: 2003-523519/49

P-PSDB; ABR62590.

New fusion protein comprising functionally linked components of an anti-CD7 antibody or its fragment and an immunotoxin, useful for treating CD7 positive acute T-cell and/or myeloid leukemias, or Graft-versus-host disease.

Claim 11: Page 10; 55pp; English.

The present sequence encodes a novel fusion protein comprising an scFv fragment of anti-CD7 monoclonal antibody TH-69 and domains II and III of Pseudomonas exotoxin A (ETA), joined via a peptide linker, and including an N-terminal 6xHis tag and C-terminal REDK sequence. This novel fusion protein binds to CD7-positive T-lymphoid cells and kills them by the induction of apoptosis. It is an example of anti-CD7 scFv immunotoxin fusion proteins of the invention. An expression vector, host cells, a method for producing the fusion protein, and a vaccine comprising the fusion protein are claimed. The fusion protein is used in the treatment of disorders involved a hyperproliferation of CD7-positive cells, especially acute T-cell and/or myeloid leukaemia, and also in the treatment or prophylaxis of Graft-versus-host disease. A claimed method of conditioning an animal (e.g. a human) to be transplanted with donor cells, tissue or organ comprises: (a) depleting the CD7-bearing cell population in the animal; (b) providing a transplant comprising isolated bone marrow and/or stem cell-enriched peripheral blood cells of the donor, where the CD7-bearing cell population in (a) and/or the transplant in (b) are treated with a T-cell depleting effective amount of the fusion protein or therapeutic composition comprising the fusion protein; and (c) introducing the transplant into the animal.

Sequence 1956 BP; 353 A; 616 C; 656 G; 331 T; 0 U; 0 Other;

Query Match 56.8%; Score 421; DB 8; Length 1956;

Best Local Similarity / 4.9%; Pred. No. 3.7e-100;  
Matches 555: Conservative 0: Mismatches 180:

OV 1 GATATTGAGTGCACAGACTACATCCTCCCTGTCTGCCTCTCTGGGAGACAGAGTCACC 60

49 GATATCCAGATGACACAGACTACATCCTCCCTGTCTGCCTCTCTGGAGACAGAGTCACC 108

[illegible]

108  
168

[illegible]

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

181 AGGTTTCAGTGGGCTGTGGAACAGATATTTCCTCACCATTAAACACCTGGAGCAG 240  
229 AGGTTTCAGTGGGCTGTGGAACAGATATTTCCTCACCATTAAACACCTGGAGCAG 288  
241 GAAGATATTGGCACTTACTTTTCCCAACAGGGCAATACGCTCCGTTGGAGCTGGTGA 300  
289 GAAGATATTGCCACTTATTATTGTGACGAGTAGCAAGCTTCCGTACACGTTGGAGGG 348  
301 GGCACCAAGCTGGAAATAAACAACGTGGTGGAGGTGGTTCGTATGTTGGTGTCTGGCGGC 360  
349 GGGACCAAGCTGGAAATAAACAACGTGGTGGAGGTGGTTCGTATGTTGGTGTCTGGCGGC 408  
361 GGGCGCTCCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 420  
409 GGCAGCTCCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 468  
421 AAGCTCGGGCTCAGTGAAGATTTCCTGCAAAAGATTCTGGCTACGCAATTCATATGCTCT 480  
469 AAGCTCGGGGGTCCCTGAAACTCTCTGTGCAGCCTCTGCACTCTTTCAGTAGCTAT 528  
481 TGGATGAACCTGGGTGAAGCAGAGGCTTGGACAGGCTTGGAGTGGATGGACGGATTTAT 540  
529 GCCATGCTTGGGTTCGCCAGACTCCAGAGAGAGGCTGGAGTGGGTCGATCCATAGT 588  
541 CCTGGAGATGAGATTCTTAACATAATGGGAAATTCAGAGGCAAGGCCATCTGACAGCA 600  
589 AGTGG---TGGTTTCACTACTATCCAGACAGTGTGAAGGCCGATTCACCATCTCCAGA 645  
601 GACAAATCTCCAGACAGCCTCATGCAAGCTCAGCAGCCTGACCTCTGTGGACTCTGGG 660  
646 GATAATGCCAGGAACATCTCTATCTGCAATGAGCAGTCTGAGGCTGTGAGGACAGCGGC 705  
661 GTCTATTCTGTGCAAGATCGGGTGTGCTAGTTATGCTATGCTATGCTAGGCTCTGAGGCTCAAGGA 720  
706 ATGATTACTGTGCAAGAGACAGGTACGGGGTA---CCTCGATGCTCGGGGGCAGGG 762  
721 ACCTCAGTCACCGTCTCCTCG 741  
763 ACCAGGTCACCGTTTCCTCG 783

RESULT 11  
ID AAA39167 standard; DNA; 1457 BP.  
XX AC AAA39167;  
XX 05-SEP-2000 (first entry)  
XX Apoptobody3sc fusion protein encoding nucleotide sequence SEQ ID NO:7.  
KW Human; Fas antigen; apoptosis; apoptobody3sc; antibody; fusion gene;  
KW medical; pharmaceutical; pharmacological; biochemical; ds.  
OS Homo sapiens.  
OS Synthetic.  
XX Key Location/Qualifiers  
FH 7..1317  
CDS /\*tag= a  
FT /product= "apoptobody3sc"  
XX  
PN JP2000102389-A.  
XX  
PD 11-APR-2000.  
XX  
PF 29-SEP-1998; 98JP-00291441.  
XX  
PR 29-SEP-1998; 98JP-00291441.  
XX (HAGI/) HAGIWARA Y.  
PA (HAGI/) HAGIWARA H.  
XX

DR WPI; 2000-332086/29.  
DR P-ESDB; AAY91026.  
XX  
PT Causing apoptosis comprises transfecting fused gene to cell and reacting  
PT cell with idiotype antibody.  
XX  
PS Disclosure; Page 7; 10pp; Japanese.  
XX  
CC The present invention describes a method for causing apoptosis in which a  
CC fused gene consisting of a gene participating to apoptosis and a gene  
CC encoding at least the variable region of anti-idiotypic antibody is  
CC transfecting to a cell to express the fused gene and then an idiotype  
CC antibody is reacted with the expressed cell. The method is useful in  
CC medical, pharmaceutical, pharmacological and biochemical fields. The  
CC present sequence encodes a fusion protein designated apoptobody3sc, which  
CC is used in the exemplification of the present invention  
XX  
SQ Sequence 1457 BP; 441 A; 309 C; 340 G; 367 T; 0 U; 0 Other;  
Query Match 56.5%; Score 418.6; DB 3; Length 1457;  
Best Local Similarity 73.6%; Pred. No. 1.5e-99;  
Matches 554; Conservative 0; Mismatches 184; Indels 15; Gaps 1;  
QY 1 GATATTCAGATGACACAGACTACATCTCCCTCTCTGCTCTCTGGGAGACAGAGTACCC 60  
Db 67 GACATCCAGATGAACCACTCTCCATCCAGTCTGTCTGCATCCCTTGGAGACACAATTACC 126  
QY 61 GTCAGTTCAGGCGCAAGTCAGACATTAGGAATATTATTAACCTTGTATCAGCAGAAACCA 120  
Db 127 ATCACTTGCATGCCAGTCAGACATTGATGTTTGGTTAAGCTGGTACCAAGCAACCA 186  
QY 121 GACGAACTGTTAAATTCCTGATCTACTACATCAGATTAAGTCTGACAGGAGTCCCATCA 180  
Db 187 GGAATATGTTCTAAACTATTGATCTATTAAGGTTTCAACTTGCACACAGGCTCCCATCA 246  
QY 181 AGGTTTCAGTGGCAGGCTGTGGAACAGATTATTCCTCACCATTAAACAACCTGGAGCAG 240  
Db 247 AGGTTTAGTGGCAGTGGATCTGGAACAGTTTCAATTAACTCAGCAGCTGAGCT 306  
QY 241 GAAGATATTGGCACTTACTTTTGGCAACAGGCAATACGCTTCCGTGAGCGTTCCGTGGA 300  
Db 307 GAAGACATTGCCACTTACTTCTCAACAGGCTCAAGGTTATCTCTCGACGCTCGGTGGA 366  
QY 301 GGCACCAAGCTGGAATAAAGCTGTGGAGGTGCTC-----TGATGT 345  
Db 367 GGCACCAAGCTGGAATAAAGCTGTGGAGGTGCTC-----TGATGT 426  
QY 346 GGTGTTCTGGCGGCGGCTCCGCTGCTGGTGGATCGAGGTCCTCAACAGCTT 405  
Db 427 GCGGTTTCAGGCGGAGTGGCTCTGGCGTGGCGGATCAGAGGTTTCCAGAGCTT 486  
QY 406 GGCAGTGGTGGTGAAGCTGGGCTCAGTGAAGATTTCTCTGCAAGATTTCTGGCTAC 465  
Db 487 GGCAGTGGTGGTGAAGCTGGGCTCAGTGAAGATTTCTCTGCAAGATTTCTGGCTAC 546  
QY 466 GCATTCAATAGCTTCTTGGATGAACCTGGGTGAAGCAGAGGCTTGGAGGCTTGGAGTGG 525  
Db 547 ACCTTTTAAACAGCTTCTGATGATGCTGGGTGAAGCAGAGGCTTGGAGTGG 606  
QY 526 ATTGGACGATTTATCTCTGAGATGAGATTTCTAATCAATGGGAAATTCAGAGGCAAG 585  
Db 607 ATTGGCGGATTTATCTCTGGAATATGATGATTTAGCTACAGCCAGAACTTTAAGGACAG 666  
QY 586 GCCATATGACAGACAGCAAAATCTCCAGCAGACCTACATGAGCTCAGAGCCTGACCC 645  
Db 667 GCCAAATGACTGCTGCTGATCCAGCAGCTGCTTACATGGAATTCAGAGGCTGACA 726  
QY 646 TCTGTGAGCTCTGGCTCTATTCTGTGCAAGATCGGGGTGTGCTAGCTTATGCTAGGAC 705  
Db 727 AATGAGGACTCTGGCTCTATTCTGTACAAAGAGGAAATATGATPACGACCCCTGGAC 786  
QY 706 TACTGGGCTCAAGGAACCTCAGTCCAGCTCC 738











CC for the depletion of B- cells and more particularly non-Hodgkin lymphoma  
CC in mammals preferably human. They can also delay the pathological  
CC conditions caused by these diseases, and can be used for detecting these  
CC diseases. The polynucleotide is used for gene therapy. The polypeptides  
CC are also used for identifying compounds modulating B-cell/T-cell mediated  
CC immune response with can in turn be used for treating cancer, its related  
CC diseases and also for inhibiting viral diseases by preventing viral  
CC infection

XX Sequence 1611 BP; 402 A; 396 C; 440 G; 373 T; 0 U; 0 Other;

```
Query Match          53.2%; Score 394.2; DB 3; Length 1611;
Best Local Similarity 73.5%; Pred. No. 3.7e-93;
Matches 564; Conservative 0; Mismatches 158; Indels 45; Gaps 3;

QY 1 GATATTTCAGATGACAGACTACATCCTCCCTGCTGCTCTCTGGGACACAGATCACC 60
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 GTCAGTTTGCAGGCGCAAGTCAG-----GACATTAGGAATTTAAACTGGTAT 108
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 152 ATCTCTGCAAGGCCAGCCAAAGTTGATTATGATGGTGATAGTTATTTGAACGGTAC 211
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 109 CAGCAGAAACAGACGGAACGTAAATTCCTGATCTACTACATCAAGATTACTGCCA 168
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 212 CAACAGATTCCAGGACAGCCACCAACTCCTCATCTATGATGCATCCAAATCTAGTTCT 271
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 169 GGAGTCCCATCAAGGTTTCAGTGGCAGTGGTCTGGAAACAGATTATTCCTCACCATTAAAC 228
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 272 GGGATCCACCCAGGTTTATGTCAGTGGTCTGGGACAGACTTCACCCCTCAACATCCAT 331
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 229 AACCTGAGCAGAGAGATATGGACCTTACTTTTGGCAACAGGGCAATAGCCTCCGTGG 288
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 332 CCTGTGGAGAAGTGGATGCTGCAACCTATCACTGTGACAAAGTACTAGGATCCGTGG 391
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 289 ACGTTCCGTTGGAGCACCAAGCTGGAATAAAGCTGGTGGAGTGGTTCTGATGGTGGT 348
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 392 ACGTTCGTTGGAGGACCAAGCTCGATCAAA-----GGTGGTGGT 433
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 349 GGTTCGCGCGCGCGCTCCGTTGGTGGATCCGAGGTCCAACTGCMAAGTCTGGA 408
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 434 GGTTCGCGCGCGCGCTCCGTTGGTGGTGGTTCACAGGTGCGAGTGCAGCAGTCTGGG 493
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 409 CCTGAGCTGGTGAAGCCTGGGGCTCAGTGAAGATTTCTGCAAGATTTCTGGCTACGCA 468
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 494 GCTGAGCTGGTGAAGCCTGGGTCTCTGATGAAGATTTCTGCAAGGCTTCTGGCTATGCA 553
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 469 TTCATAGCTTTTGGATGAACCTGGTGAAGCAGAGGCTCGACAGGCTCTTGAGTGGATT 528
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 554 TTCAGTAGCTACTGGATGAACCTGGTGAAGCAGAGGCTCGACAGGCTCTTGAGTGGATT 613
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 529 GGCGGATTTTATCCTGGAGATGGAGATTCTAACTACAATGGAAATTCGAGGGCAAGGCC 588
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 614 GGACAGATTTGGCTGGAGATGGTGATCTAACTACAATGGAAATTCAGGGTAAAGCC 673
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 589 ATACTGACAGACAGAAATCCTCAGACAGCCTTACATGAGCTCAGACGCTGACCTCT 648
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 674 ACTCTGACTGACAGCAATCCTCAGACAGCCTTACATGCAACTCAGACGCTAGCATCT 733
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 649 GTGACTCTGGGCTTATTTCTGTGCAAGATCG-----GGGTTGCTACGT 693
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 734 GAGGACTCTGGGCTTATTTCTGTGCAAGCAGGAGACTAGCAGGTAGGCCGCTTATTAC 793
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 694 TATGCTATGACTACTGGGGTCAAGGAACTCAGTACCCGCTCTCCTC 740
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 794 TATGCTATGACTACTGGGGCAAGGACCAACGCTCACCCTCTCCTC 840
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: April 26, 2004, 16:01:08  
Job time : 399 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2004, 15:43:37 ; Search time 3232 Seconds  
(without alignments)  
9937.249 Million cell updates/sec

Title: US-10-620-049-24

Perfect score: 741

Sequence: 1 gatattcagatgacacagac.....ctctacgacgctctctcgcg 741

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Result No.	Score	Query Match %	Length	DB ID	Description
1	499.4	67.4	1803	6	BD237850 Anti-CD3
2	499.4	67.4	1940	6	AX306601 Sequence
3	499.4	67.4	1943	6	AX306600 Sequence
4	482.8	65.2	2952	6	AX306590 Sequence
5	479.2	64.7	916	6	AR116105 Sequence
6	479.2	64.7	916	6	I45604 Sequence 29
7	479.2	64.7	916	6	AR256099 Sequence
8	479.2	64.7	916	6	AR401482 Sequence
9	456.2	61.6	748	12	AF326551 Synthetic
10	431.8	58.3	2691	6	AX306662 Sequence
11	421	56.8	1953	6	AX797328 Sequence
12	421	56.8	1956	6	AX797327 Sequence
13	418.6	56.5	1457	6	E35160 Method for
14	406	54.8	1545	6	AX398258 Sequence
15	404.2	54.5	1698	6	AX011206 Sequence
16	404.2	54.5	1698	6	BD206134 Polyvalen
17	396	53.4	888	12	AF132308 Synthetic
18	395.8	53.4	1817	6	AX739885 Sequence
19	394.2	53.2	1611	6	AX014270 Sequence
20	394.2	53.2	1611	6	AX839766 Sequence
21	394.2	53.2	1611	6	BD205046 CD19XCD3-
22	393	53.0	902	12	XXU49832 Synthetic s
23	385.6	52.0	1637	6	AX006750 Sequence
24	385.6	52.0	1637	6	BD222155 Immunolog
25	381.8	51.5	819	6	AI8692 Synthetic n
26	374.8	50.6	1817	6	AX739887 Sequence
27	371.2	50.1	1612	6	AX057943 Sequence
28	368	49.7	786	6	AX798501 Sequence
29	368	49.7	993	6	AX798489 Sequence
30	368	49.7	1017	6	AX798479 Sequence
31	363.8	49.1	2871	6	AX023363 Sequence
32	363.8	49.1	2871	6	BD222937 Heteromin
33	361.8	48.8	1653	6	AX011208 Sequence
34	361.8	48.8	1653	6	BD206135 Polyvalen
35	358.6	48.4	877	6	E31225 Device for
36	358.6	48.4	925	6	E30616 Antibody an
37	356.8	48.2	840	12	AF329458 Synthetic
38	356.4	48.1	1299	6	AR060677 Sequence
39	356.4	48.1	1299	6	AR074434 Sequence
40	356.4	48.1	1320	6	AR060674 Sequence
41	356.4	48.1	1320	6	AR074431 Sequence
42	356.4	48.1	6727	6	AR060673 Sequence
43	356.4	48.1	6727	6	AR074430 Sequence
44	356.4	48.1	6799	6	AR060676 Sequence
45	356.4	48.1	6799	6	AR074433 Sequence

ALIGNMENTS

RESULT 1  
LOCUS BD237850 1803 bp DNA linear PAT 17-JUL-2003  
DEFINITION Anti-CD3 immunotoxins and therapeutic uses therefor.  
ACCESSION BD237850  
VERSION BD237850.1 GI:33047620  
KEYWORDS JP 2002534441-A/1.  
SOURCE synthetic construct  
ORGANISM artificial sequences.  
REFERENCE 1 (bases 1 to 1803)  
AUTHORS Digan,M.E., Lake,P. and Wright,R.M.  
TITLE Anti-CD3 immunotoxins and therapeutic uses therefor  
JOURNAL Patent: JP 2002534441-A 1 15-OCT-2002;  
NOVARTIS AG

Pred. No. is the number of results predicted by chance to have a

```

COMMENT      OS      Artificial Sequence
PN      JP 2002534441-A/1
PD      15-OCT-2002
PR      13-JAN-2000 JP 2000593098
PF      15-JAN-1999 US 09/232445,25-JAN-1999 US 09/236968 PR
PI      07-OCT-1999 US 09/414134
PI      MARY ELLEN DIGAN, PHILIP LAKE, RICHARD MICHAEL WRIGHT PC
A61K39/395, A61K31/137, A61K31/27, A61K31/343, A61K31/42, A61K31/PC
436, A61K31/52, A61K31/573, A61K31/664, A61K31/7056, A61K38/00, A61K39/104 PC
PC      A61K31/573, A61K31/664, A61K31/7056, A61K38/00, A61K39/104 PC
PC      C07K16/28, C07K19/00, C12N15/09//C07D261/18, C07D307/88 PC
PC      C07D473/38, C07D498/18,
PC      C12N15/00, A61K37/02
CC      Description of Artificial Sequence: scFv (UCHT-1)-PE28 PH Key
FT      CDS      Location/Qualifiers
              1..1803
              /organism="synthetic construct"
              /mol_type="genomic DNA"
              /db_xref="taxon:32630"
FEATURES
    source
    1..1803
    /organism="synthetic construct"
    /mol_type="genomic DNA"
    /db_xref="taxon:32630"
ORIGIN
Query Match      67.4%; Score 499.4; DB 6; Length 1803;
Best Local Similarity 81.4%; Pred. No. 2.2e-136;
Matches 610; Conservative 0; Mismatches 121; Indels 18; Gaps 2;
QY      1 GATATTGAGTACACAGTACATCTCTCCCTGTCCTCTCTGGGAGACAGATCACC 60
DB      7 GACATCCAGATGACCCAGACACCTCTCTCCCTGTCCTCTCTGGGAGACAGATCACC 66
QY      61 GTCAGTTGACGGCAAGTCAGGACATAGGAAATATTATTAACCTGGTATACAGCAAAACCA 120
DB      67 ATCAGTTGACGGCAAGTCAGGACATAGGAAATATTATTAACCTGGTATACAGCAAAACCA 126
QY      121 GACGGAAGCTGTAATTCCTGATCTCTACATCAATCAAGATTAAGTCCAGAGTCCCATCA 180
DB      127 GATGGAAGCTGTAATTCCTGATCTCTACATCAATCAAGATTAAGTCCAGAGTCCCATCA 186
QY      181 AGGTTGAGTGGCAGTGGGTCTGGAACAGATTAATTCCTCCACCAATTAACACCTGGAGCAG 240
DB      187 AAGTTGAGTGGCAGTGGGTCTGGAACAGATTAATTCCTCCACCAATTAACACCTGGAGCAG 246
QY      241 GAAGATATTGGCACTTACTTTTGGCAACAGGGCAATAGCGCTCGGTGGAGTTCGGTGGGA 300
DB      247 GAGGATATTGGCACTTACTTTTGGCAACAGGGCAATAGCGCTCGGTGGAGTTCGGTGGGA 306
QY      301 GGCACCAAGCTGGAAATAAAACGTTGGTGGAGTGGTTCCTGATGGTGGTTCCTGGCGGC 360
DB      307 GGCACCAAGCTGGAAATAAAACGTTGGTGGAGTGGTTCCTGATGGTGGTTCCTGGCGGC 366
QY      361 GCGCGCTCGGTGGTGGTGGATCCGAGTCCAACTGCAACAGTCTGACCTGAGCTGGTG 420
DB      367 AGC-----GGTGGCGGATCTGAGGTGCGAGTCCAGCAGTCTGGACCTGAGCTGGTG 417
QY      421 AAGCTGGGGCTCAGTGAAGATTTCTGCAAGATTTCTGCTACGCAATTCATATAGCTCT 480
DB      418 AAGCTGGAGCTTCAATGAAGATATCTGCAAGCTTCTGTTACTCATTCAGTGGCTAC 477
QY      481 TGGATGAAGCTGGGTGAAGAGCGCTGGACAGGGTCTTGAGTGGATTTGACCGGATTTAT 540
DB      478 ACCATGAAGCTGGGTGAAGAGCGCTGGACAGGGTCTTGAGTGGATTTGACCGGATTTAT 537
QY      541 CCTGGAGATGGAGATTTCTAATCAATGGGAAATTCGAGGGCAAGGCCATCTGACAGCA 600
DB      538 CTTTACAAAGGTGTAGTACTCTACAAACGAGTTCAAGCAAGGCCACATTAACCTGTA 597
QY      601 GACAAATCCTCCAGCAGAGCTTACATGAGCTAGCAGCTGAGCTCTGTTGGACTCTGGC 660
DB      598 GACAAGTCAATCAGCAGAGCTTACATGAGAACTCTCAGTCTGACATCTGAGGACTCTGCA 657
QY      661 GTCTATTCTGCAAGATCGGGTGTGCTAGCTTATGCT-----ATGGACTACTGG 711

```

```

DB      658 GTCTATTACTGTGCAAGATCGGGTACTACCGTGATAGTACTGGTACTCGATGTCCTGG 717
QY      712 GGTCAAGAACTCAGTCACTCAGTCACCGTCTCCCTC 740
DB      718 GGCAGGAGGACCAAGGTCAACCGTCTCCCTC 746
RESULT 2
LOCUS      AX306601
DEFINITION Sequence 41 from Patent WO0187982.
ACCESSION AX306601
VERSION    AX306601.1 GI:17645800
KEYWORDS   synthetic construct
SOURCE     synthetic construct
ORGANISM   artificial sequences.
REFERENCE  1
AUTHORS    Neville,D.M., Thompson,J.T., Hu,H., Woo,J.H., Ma,S., Hexham,J.M.
            and Digan,M.E
TITLE      Immunotoxin fusion proteins and means for expression thereof
JOURNAL    Patent: WO 0187982-A 41 22-NOV-2001;
            THE DEPARTMENT OF HEALTH & HUMAN SERVICES (US)
FEATURES   Location/Qualifiers
            1..1940
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="Note = synthetic construct"
ORIGIN
Query Match      67.4%; Score 499.4; DB 6; Length 1940;
Best Local Similarity 81.4%; Pred. No. 2.2e-136;
Matches 610; Conservative 0; Mismatches 121; Indels 18; Gaps 2;
QY      1 GATATTGAGTACACAGTACATCTCTCCCTGTCCTCTCTGGGAGACAGATCACC 60
DB      1188 GACATCCAGATGACCCAGACACCTCTCTCCCTGTCCTCTCTGGGAGACAGATCACC 1247
QY      61 GTCAGTTGACGGCAAGTCAGGACATTAGGAAATATTATTAACCTGGTATACAGCAAAACCA 120
DB      1248 ATCAGTTGACGGCAAGTCAGGACATTAGGAAATATTATTAACCTGGTATACAGCAAAACCA 1307
QY      121 GACGGAAGCTGTAATTCCTGATCTCTACATCAATCAAGATTAAGTCCAGAGTCCCATCA 180
DB      1308 GATGGAAGCTGTAATTCCTGATCTCTACATCAATCAAGATTAAGTCCAGAGTCCCATCA 1367
QY      181 AGGTTGAGTGGCAGTGGGTCTGGAACAGATTAATTCCTCCACCAATTAACACCTGGAGCAG 240
DB      1368 AAGTTGAGTGGCAGTGGGTCTGGAACAGATTAATTCCTCCACCAATTAACACCTGGAGCAG 1427
QY      241 GAAGATATTGGCACTTACTTTTGGCAACAGGGCAATAGCGCTCGGTGGAGTTCGGTGGGA 300
DB      1428 GAGGATATTGGCACTTACTTTTGGCAACAGGGCAATAGCGCTCGGTGGAGTTCGGTGGGA 1487
QY      301 GGCACCAAGCTGGAAATAAAACGTTGGTGGAGTGGTTCCTGATGGTGGTTCCTGGCGGC 360
DB      1488 GGCACCAAGCTGGAAATAAAACGTTGGTGGAGTGGTTCCTGATGGTGGTTCCTGGCGGC 1547
QY      361 GCGCGCTCGGTGGTGGTGGATCCGAGTCCAACTGCAACAGTCTGGAACCTGAGCTGGTG 420
DB      1548 AGC-----GGTGGCGGATCTGAGGTGCGAGTCCAGCAGTCTGGACCTGAGCTGGTG 1598
QY      421 AAGCTGGGGCTCAGTGAAGATTTCTGCAAGATTTCTGCTACGCAATTCATATAGCTCT 480
DB      1599 AAGCTGGAGCTTCAATGAAGATATCTTGAAGGCTTCTGTTACTCATTCAGTGGCTAC 1658
QY      481 TGGATGAAGCTGGGTGAAGAGCGCTGGACAGGGTCTTGAGTGGATTTGACCGGATTTAT 540
DB      1659 ACCATGAAGCTGGGTGAAGAGCGCTGGACAGGGTCTTGAGTGGATTTGACCGGATTTAT 1718
QY      541 CTTGAGATGGAGATTTCTAATCAATGGGAAATTCGAGGGCAAGGCCATCTGACAGCA 600

```



Db 1735 GGCACCAAGCTGGAGATATAAAGGAGGCGGAGGACG-----CGGAGGC 1776  
QY 361 GCGGCTCCGGTGGTGGATCGAGGTCCAACTGCAACAGTCTGGACCTGAGCTGGTG 420  
Db 1777 GGTGGCTCGGAGGCGGAGGCTCGAGGTGCGAGTCCAGCAGTCTGGACCTGAGCTGGTG 1836  
QY 421 AAGCCTGGGCGCTCAGTGAAGATTCTCGAAGATTCTGGACGAGCTCAATCAATAGCTCT 480  
Db 1837 AAGCCTGGAGCTTCAATGAAGATATCTCGCAAGGCTTCTGGTACTCAITCACTGGGTAC 1896  
QY 481 TGGATGAAGTGGGGAAGCAGAGCGCTGGACAGGCTCTTGAGTGGATTGGACGGATTAT 540  
Db 1897 ACCATGAAGTGGGGAAGCAGAGCTATGGAAGAACCTTGAGTGGATTGGACCTTATTAAT 1956  
QY 541 CTTGGAGATGGAGATTCTAATCAATGGAATTCGAGGCAAGGCCATPACTGACAGCA 600  
Db 1957 CTTTACAAAGGTGTAGTACTCAACACCAAGATTCGAGGCAAGGCCATTAATCTGTA 2016  
QY 601 GACAAATCTCCAGCAGCCTACATGAGCTCAGAGCTGAGCCTCTGTGGACTCTGG 660  
Db 2017 GACAAATCTCCAGCAGCCTACATGAGCTCAGAGCTGAGCCTCTGTGGACTCTGG 2076  
QY 661 GTCTATTCTGTGCAAGATCGGCTGTGCTAGTATGCTATGGAGCTACTGGGCTCAAGGA 720  
Db 2077 GTCTATTACTGTGCAAGATCGGCTGTGCTAGTATGCTATGGAGCTACTGGGCTCAAGGA 720  
QY 721 ACTCAGTCAACC 732  
Db 2137 GCGCAGGACACC 2148

RESULT 5  
LOCUS AR116105 916 bp DNA linear PAT 16-MAY-2001  
DEFINITION Sequence 29 from patent US 6132992.  
ACCESSION AR116105  
VERSION AR116105.1 GI:14096427  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 916)  
AUTHORS Ledbetter,J.A., Gilliland,L.K., Hayden,M.S., Linsley,P.S.,  
Bajorath,J. and Fell,H.Perry.  
TITLE Expression vectors encoding bispecific fusion proteins and methods  
of producing biologically active bispecific fusion proteins in a  
mammalian cell  
JOURNAL Patent: US 6132992-A 29 17-OCT-2000;  
FEATURES Location/Qualifiers  
source 1..916  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 64.7%; Score 479.2; DB 6; Length 916;  
Best Local Similarity 79.4%; Pred. No. 2e-130;  
Matches 581; Conservative 0; Mismatches 148; Indels 3; Gaps 1;

QY 1 GATATTTCAGATGACACAGACTACATCTCCCTGTCTGCTCTCTCTGGGAGACAGAGTCAACC 60  
Db 76 GACATCCAGATGACACAGACTACATCTCCCTGTCTGCTCTCTCTGGGAGACAGAGTCAACC 135  
QY 61 GTCAGTTTCAGGCGAGTTCAGGACATAGGATTTAACTGGTATCAGCAGAAACCA 120  
Db 136 ATCAGTTTCAGGCGAGTTCAGGACATAGGATTTAACTGGTATCAGCAGAAACCA 195  
QY 121 GACGGAATCTGTAAATTCCTGATCTACTACATCAAGATTACTGCCAGGAGTCCCATCA 180  
Db 196 GATGGAATCTGTAAATTCCTGATCTACTACATCAAGATTACTGCCAGGAGTCCCATCA 255  
QY 181 AGGTTTCAGTGGAGTGGTTCGGAACAGATTATTCCTCACCATTAAACCTGGAGCA 240

Db 256 AGGTTTCAGTGGAGTGGTCTGGAAACAGATTATTTCTCTACCAATTGCCAACCTGCAACCA 315  
QY 241 GAAGATATTGGCACTTACTTTTGCCAAACAGGCGCAATAGCCTCCGTTGACCTTCGGTGA 300  
Db 316 GAAGATATTGGCACTTACTTTTGCCAAACAGGCGTAATAGCCTTCGTTGGACCTTCGGTGA 375  
QY 301 GGCACCAAGCTGGAAATATAAAGCTGGTGGAGTGGTCTGATGGTGGTGGTCTTCGGGGC 360  
Db 376 GGCACCAACCTGGTAACCAAAACGGGAGCTCGGTGGCGGTGGCTCGGGCGGTGGGTGTCG 435  
QY 361 GCGCGCTCCGGTGGTGGTGGATCCGAGGTCCAACTGCAACAGTCTGGAACCTGAGTGGTG 420  
Db 436 GGTGG---CGCGGATCTATCGATGAGGTCCAGCTGCAACAGTCTGGAACCTGAGTGGTG 492  
QY 421 AAGCCTGGGCGCTCAGTGAAGATTCTCTGAAAAGATTCTGGCTAGCGATTCAATAGCTCT 480  
Db 493 AAGCCTGGAGCTTCAATGACAAATGCTCTGCAAGGCTCTGTTACTCAITCACTGGGTAC 552  
QY 481 TGGATGAAGTGGTGGTGGTGGATCCGAGGTCCGACAGGCTCGACAGGCTCTTGAGTGGATTGGACGGATTAT 540  
Db 553 ATCTGTGAAGTGGTGGTGGATCCGAGGCTCGACAGGCTTGGAGTGGATTGGACTTATTAAT 612  
QY 541 CTTGGAGATGGAGATTCTAATCAATGGGAAATTCGAGGCGCAAGGCCATPACTGACAGCA 600  
Db 613 CCATACAAAGGTCTTACTACCTACCAACCAAGAAATTCAGAGGCGCAAGGCCACATTAATCTGA 672  
QY 601 GACAAATCTCCAGCAGCCTACATGAGCTCAGAGCTCAGCAGCCTGAGCTCTGTGGACTCTGG 660  
Db 673 GACAAATCTCCAGCAGCCTACATGAGCTCCTCAGTCTGACATCTGAAAGACTCTGCA 732  
QY 661 GTCTATTCTGTGCAAGATCGGCTGTGCTAGTATGCTATGGACTACTGGGCTCAAGGA 720  
Db 733 GTCTATTACTGTGCAAGATCTGGTACTATGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 792  
QY 721 ACTCAGTCAACC 732  
Db 793 GCGCAGGACACC 804

RESULT 6  
LOCUS 145604 916 bp DNA linear PAT 07-OCT-1997  
DEFINITION Sequence 29 from patent US 5637481.  
ACCESSION 145604  
VERSION 145604.1 GI:2469706  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 916)  
AUTHORS Ledbetter,J.A., Gilliland,L.K., Hayden,M.S., Linsley,P.S.,  
Bajorath,J. and Fell,H.Perry.  
TITLE Expression vectors encoding bispecific fusion proteins and methods  
of producing biologically active bispecific fusion proteins in a  
mammalian cell  
JOURNAL Patent: US 5637481-A 29 10-JUN-1997;  
FEATURES Location/Qualifiers  
source 1..916  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 64.7%; Score 479.2; DB 6; Length 916;  
Best Local Similarity 79.4%; Pred. No. 2e-130;  
Matches 581; Conservative 0; Mismatches 148; Indels 3; Gaps 1;

QY 1 GATATTTCAGATGACACAGACTACATCTCCCTGTCTGCTCTCTCTGGGAGACAGAGTCAACC 60  
Db 76 GACATCCAGATGACACAGACTACATCTCCCTGTCTGCTCTCTCTGGGAGACAGAGTCAACC 135  
QY 61 GTCAGTTTCAGGCGAGTTCAGGACATAGGATTTAACTGGTATCAGCAGAAACCA 120  
Db 136 ATCAGTTTCAGGCGAGTTCAGGACATAGGATTTAACTGGTATCAGCAGAAACCA 195



QY 121 GAGGAACTGTTAAATTCCTGATCTACTACACATCAAGATTACTGCCAGGAGTCCCATCA 180  
Db 196 GATGGAACTGTTAAATTCCTGATCTACTACACATCAAGATTACTGCCAGGAGTCCCATCA 255  
QY 181 AGTTTCAGTGGCAGTGGTCTGGACAGATATTCCTCAGCATTAACACCTGGAGCAG 240  
Db 256 AGTTTCAGTGGCAGTGGTCTGGACAGATATTCCTCAGCATTAACACCTGGAGCAG 315  
QY 241 GAAGATATGGCACTTACTTTTGGCAACAGGCAATACGCTCCGTGGAGTTCGGTGGG 300  
Db 316 GAAGATATGGCACTTACTTTTGGCAACAGGCAATACGCTCCGTGGAGTTCGGTGGG 375  
QY 301 GGCAACAGCTGGAAATTAACAGTGGTGGAGTGGTCTGATGGTGGTGGTGGTGGG 360  
Db 376 GGCAACAGCTGGAAATTAACAGTGGTGGAGTGGTCTGATGGTGGTGGTGGTGGG 420  
QY 361 GGCGGCTCCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 435  
Db 436 GGTGG---GGCGGATCTATCGATGAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG 492  
QY 421 AAGCTTGGGCTCAGTGAAGATTTCTGCAAGATTTCTGGTACGCAATTAAGTCT 480  
Db 493 AAGCTTGGGCTCAGTGAAGATTTCTGCAAGATTTCTGGTACGCAATTAAGTCT 552  
QY 481 TGGATGAAGTGGTGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 540  
Db 553 ATCGTGAAGTGGTGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 612  
QY 541 CCTGGAGATGAGATTTCTAACTAATGGGAAATTCGAGGGCAAGGCCATTAAGTGA 600  
Db 613 CCATACAAAGTCTTACTACCTACCAAGGAAATTCGAGGGCAAGGCCATTAAGTGA 672  
QY 601 GACAAATCCCTCCAGCAGCAGCTTACATGAGCTCAGCAGCTCAGCAGCTCAGCAG 660  
Db 673 GACAAATCCCTCCAGCAGCAGCTTACATGAGCTCAGCAGCTCAGCAGCTCAGCAG 732  
QY 661 GTCTATTCTGTGCAAGATCGGGTGTCTAGTATGCTATGAGTCTGAGTCTGAGTCA 720  
Db 733 GTCTATTACTGTGCAAGATCGGGTGTCTAGTATGCTATGAGTCTGAGTCTGAGTCA 792  
QY 721 ACCTCAGTCACC 732  
Db 793 GCGCAGGGGACC 804

RESULT 7  
AR256099 AR256099 916 bp DNA linear PAT 20-DEC-2002  
LOCUS Sequence 29 from patent US 6482919.  
DEFINITION AR256099  
ACCESSION AR256099.1 GI:27305471  
VERSION  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 916)  
AUTHORS Ledbetter,J.A., Hayden,M.S., Linsley,P.S., Bajorath,J., Fell,H.P.  
and Gilliland,L.K.  
TITLE Expression vectors encoding bispecific fusion proteins and methods of producing biologically active bispecific fusion proteins in a mammalian cell  
JOURNAL Patent: US 6482919-A 29 19-NOV-2002;  
FEATURES Location/Qualifiers  
source 1..916  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 64.7%; Score 479.2; DB 6; Length 916;  
Best Local Similarity 79.4%; Pred. No. 2e-130;  
Matches 581; Conservative 0; Mismatches 149; Indels 3; Gaps 1;

QY 1 GATATTGAGATGACACAGCTACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60  
Db 76 GACATCCAGATGACACAGCTACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 135  
QY 61 GTCACTTCAGGCGCAAGTACAGCAATTAAGAAATTAATTAATTAATTAATTAATTAAT 120  
Db 136 ATCACTTCAGGCGCAAGTACAGCAATTAAGAAATTAATTAATTAATTAATTAATTAAT 195  
QY 121 GACGGAATGTTAAATTCCTGATCTACTACATCAAGATTACTGCCAGGAGTCCCATCA 180  
Db 196 GATGGAATGTTAAATTCCTGATCTACTACATCAAGATTACTGCCAGGAGTCCCATCA 255  
QY 181 AGTTTCAGTGGCAGTGGTCTGGACAGATATTCCTCAGCATTAACACCTGGAGCAG 240  
Db 256 AGTTTCAGTGGCAGTGGTCTGGACAGATATTCCTCAGCATTAACACCTGGAGCAG 315  
QY 241 GAAGATATGGCACTTACTTTTGGCAACAGGCAATACGCTCCGTGGAGTTCGGTGGG 300  
Db 316 GAAGATATGGCACTTACTTTTGGCAACAGGCAATACGCTCCGTGGAGTTCGGTGGG 375  
QY 301 GGCAACAGCTGGAAATTAACAGTGGTGGAGTGGTCTGATGGTGGTGGTGGTGGG 360  
Db 376 GGCAACAGCTGGAAATTAACAGTGGTGGAGTGGTCTGATGGTGGTGGTGGTGGG 435  
QY 361 GGCGGCTCCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 420  
Db 436 GGTGG---GGCGGATCTATCGATGAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG 492  
QY 421 AAGCTTGGGCTCAGTGAAGATTTCTGCAAGATTTCTGGTACGCAATTAAGTCT 480  
Db 493 AAGCTTGGGCTCAGTGAAGATTTCTGCAAGATTTCTGGTACGCAATTAAGTCT 552  
QY 481 TGGATGAAGTGGTGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 540  
Db 553 ATCGTGAAGTGGTGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 612  
QY 541 CCTGGAGATGAGATTTCTAACTAATGGGAAATTCGAGGGCAAGGCCATTAAGTGA 600  
Db 613 CCATACAAAGTCTTACTACCTACCAAGGAAATTCGAGGGCAAGGCCATTAAGTGA 672  
QY 601 GACAAATCCCTCCAGCAGCAGCTTACATGAGCTCAGCAGCTCAGCAGCTCAGCAG 660  
Db 673 GACAAATCCCTCCAGCAGCAGCTTACATGAGCTCAGCAGCTCAGCAGCTCAGCAG 732  
QY 661 GTCTATTCTGTGCAAGATCGGGTGTCTAGTATGCTATGAGTCTGAGTCTGAGTCA 720  
Db 733 GTCTATTACTGTGCAAGATCGGGTGTCTAGTATGCTATGAGTCTGAGTCTGAGTCA 792  
QY 721 ACCTCAGTCACC 732  
Db 793 GCGCAGGGGACC 804

RESULT 8  
AR401482 AR401482 916 bp DNA linear PAT 18-DEC-2003  
LOCUS Sequence 29 from patent US 6623940.  
DEFINITION AR401482  
ACCESSION AR401482  
VERSION AR401482.1 GI:40148818  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 916)  
AUTHORS Ledbetter,J.A. and Linsley,P.S.  
TITLE Expression vectors encoding bispecific fusion proteins and methods of producing biologically active bispecific fusion proteins in a mammalian cell  
JOURNAL Patent: US 6623940-A 29 23-SEP-2003;  
FEATURES Location/Qualifiers  
source 1..916  
/organism="unknown"  
/mol\_type="genomic DNA"

<b>ORIGIN</b>					
Query Match            64.7%; Score 479.12; DB 6; Length 916;					
Best Local Similarity   79.4%; Pred. No. 2e-130;					
Matches 581; Conservative   0; Mismatches 148; Indels   3; Gaps   1;					
Qy	1	GATATTGAGTACACAGACTACATCCTCCCTGTCTGCCTCTCTGGAGACAGAGTCACC	60		
Dd	76	GACATCCAGATGACACAGACTACATCCTCCCTGTCTGCCTCTCTGGAGACAGAGTCACC	135		
Qy	61	GTCAAGTTGCAGGCGAAGTCAGGACATTAGGAATTAATTAAACTGGTATCAGCAGAACA	120		
Dd	136	ATCAGTTGCAGGCGAAGTCAGGACATTCGCAATTAATTAACTGGTATCAGCAGAACA	195		
Qy	121	GACGGAACTGTTAAATTCCTGATCTACTACACATCAAGATTAATGCCAGAGTCCCATCA	180		
Dd	196	GATGGAACTGTTAAACTCCTGATCTACTACACATCAAGATTAATCAGAGTCCCATCA	255		
Qy	181	AGGTTCAAGTCGCGAGTGCGTCTGGAACAGATTAATTCCTCCCTCACCAATAAACCTGGAGCAG	240		
Dd	256	AGGTTCAAGTCGCGAGTGCGTCTGGAACAGATTAATTCCTCCCAATGCCAACCTGCACA	315		
Qy	241	GAAGATATTGGCACTTACTTTTTGGCCAAACAGGGCAAATACGCTCCGTGGACGTTGGTGG	300		
Dd	316	GAAGATATTGGCACTTACTTTTTGGCCAAACAGGGTAATACGCTCCGTGGACGTTGGTGG	375		
Qy	301	GGCACCAAGCTGGAAATAAACAAGTGGTGGAGGTGGTTCTGATGGTGGTGGTCTGGCGGC	360		
Dd	376	GGCACCAAACTGGTAAACAAACGGAGACTCGTGGCGGTGGTCTGGCGGTGGTGGTGG	435		
Qy	361	GGCGGCTCCGGTGGTGGTGGATCCGAGGTCCAACTGCACAGTCTGCACCTGAGCTGGTG	420		
Dd	436	GGTGG---CGGCGGATCTATCGATGAGGTCCAGCTGCAACAGTCTGGAACCTGAGTGGTG	492		
Qy	421	AAGCCTGGGGCTCAGTGAAGATTTCCTGCAAGATTCTGGCTACGCAATTAATAGTCT	480		
Dd	493	AAGCCTGGAGGCTTCAAAGCAATGCTCTGCAAGGCCCTCTGGTACTCATTCACCTGGGTAC	552		
Qy	481	TGGATGAACCTGGGTGAAGCAGAGCGCTGGACAGGGTCTTGAGTGGATTGGACGGATTAT	540		
Dd	553	ATCGTGAACCTGGCTGAAGCAGAGCCATGGAAGAACCTTGAGTGGATTGGACTTATTAT	612		
Qy	541	CCTGGAGATGAGATTCTTAAGTACATAGGGGAAATTCAGAGGCAAGGCCATCTGACAGCA	600		
Dd	613	CCATACAAAGGCTTACTACTACACACAGAAAATTCAGGGCAAGGCCACATTAACCTGTA	672		
Qy	601	GACAAATCCTCCAGCAGACCTACATGAGCTCAGCAGCTGAGCCTCTGTGGACTCTGGC	660		
Dd	673	GACAAGTCATCCAGCAGACCTACATGAGGCTCTCAGTCTGACATCTGAGACTCTGCA	732		
Qy	661	GTCTATTCTGTGCAAGATCGGGTTCCTACGTTATGCTATGGACTACTGGGGTCAAGGA	720		
Dd	733	GTCTATTACTGTGCAAGATCTGGGTACTATGGTACTCGGACTGGTACTTCGATGTCGG	792		
Qy	721	ACCTCAGTCAAC 732			
Dd	793	GGCGCAGGACC 804			
RESULT 9					
AF536561					
LOCUS	748 bp mRNA linear SYN 01-OCT-2003				
DEFINITION	Synthetic construct Mus musculus single-chain Fv antibody Klon				
ACCESSION	147.2 mRNA, partial cds.				
VERSION	AF536561				
KEYWORDS	AF536561.1 GI:33330964				
SOURCE	synthetic construct				
ORGANISM	synthetic construct				
REFERENCE	artificial sequences.				
AUTHORS	1 (bases 1 to 748)				
TITLE	Sabarth,N., Eumarn,D. and Meyer,T.F. Murine scFv-antibody against Helicobacter pylori surface protein				



Db 469 AAGCCTGGGGGTCCTGAAACTCTCTGTGAGCCTCTGAGTCACTTTCACTAGTAT 528  
Qy 481 TGGATGAATGGGTAAGACAGAGCCCTGGACAGGGTCTTGAGTGGATGAGACGGATTAT 540  
Db 529 GCCATGTTCTGGGTTGGCCAGACTCCAGAGAGAGGCTGGAGTGGGTGCGATCCATTAGT 588  
Qy 541 CCTGGAGATGGAGATTCTAACTACAATGGGAATTCGAGGGCAAGGCCATCTACGACGA 600  
Db 589 AGTGG---TGGTTTCACTACTATCCAGACAGTGTGAAGGCCGATTCACCATCTCCAGA 645  
Qy 601 GACAAATCTCCACACAGCCATACATGAGCTCAGCAGCTGAGCCTCTGTGGACTCTGGG 660  
Db 646 GATAATGCCAGGAACATCCTGTATCTGCAATAGCAGTCTGAGTCTGAGGACACGGCC 705  
Qy 661 GTCTATTCTGTGAAGATCGGGTCTGTAGTTATGCTATGAGTACTGAGGCTCAGGA 720  
Db 706 ATGATTACTGTGAAGAGAGAGATAGGGGGTA---CCTCGATGTTGGGGGCGAGGG 762  
Qy 721 ACCTCAGTCACCGTCTCCTCG 741  
Db 763 ACCACGGTCACCGTTTCCTCG 783

RESULT 12  
AX797327  
LOCUS AX797327 1956 bp DNA linear PAT 04-OCT-2003  
DEFINITION Sequence 4 from Patent WO03051926.  
ACCESSION AX797327  
VERSION AX797327.1 GI:37517969  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM artificial sequences.  
REFERENCE 1  
AUTHORS Fey, G.H., Gramatzki, M. and Peipp, M.  
TITLE Anti-cd7 immunotoxin as fusion protein  
JOURNAL Patent: WO 03051926-A 4 26-JUN-2003;  
Friedrich-Alexander-Universitaet Erlangen-Nuernberg (DE)  
FEATURES  
source  
1. 1956  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"

ORIGIN  
Query Match 56.8%; Score 421; DB 6; Length 1956;  
Best Local Similarity 74.9%; Pred. No. 3.7e-113;  
Matches 555; Conservative 0; Mismatches 180; Indels 6; Gaps 2;

Qy 1 GATATTCAGTACACAGACTACATCCCTCCTGCTGCTCTCTGGGAGACAGAGTCACC 60  
Db 49 GATATCCAGATGACAGACTACATCCCTCCTGCTGCTCTCTGGGAGACAGAGTCACC 108  
Qy 61 GTCAGTTGAGGGCAAGTCAGACATTTAGGAATTTAACTGGTATCAGCAGAAACCA 120  
Db 109 ATCAGTTGAGTCAAGTCAGGCAATTTAGCAATTTAACTGGTATCAGCAGAAACCA 168  
Qy 121 GACGAGCTGTTAAATCTCTGATCTACTACATCAAGATTACTCCAGGAGTCCATCA 190  
Db 169 GATGGAGCTGTTAAATCTCTGATCTATTACATCAAGTTACTCAGGAGTCCATCA 228  
Qy 181 AGTTTCAGTGGCAGTGGGTCTGGAACAGATTATTCCTCCACCATPAAACCTGGAGAG 240  
Db 229 AGTTTCAGTGGCAGTGGGTCTGGAACAGATTATTCCTCCACCATCAGCAACCTGAACT 288  
Qy 241 GAGATATTGGCACTTACTTTTGGCAACAGGCAATAGCCTCCGTGGAGCTGGTGGGA 300  
Db 289 GAGATATTGGCACTTATTATTGTGACAGATATAGCAAGCTTCCTGATCAGCTTCGAGGG 348  
Qy 301 GGCACCAAGCTGGAATAAAGCTGGTGGAGTGGTCTTGATGGTGGTCTGGCGGC 360  
Db 349 GGGACCAAGCTGGAATAAAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 408

Qy 361 GCGGCTCGGCTGGTGGTGGATCCAGTCCAACTGCAACAGTCTGAGCTGAGCTGGTG 420  
Db 409 GCGGCTCGGCTGGTGGTGGATCCAGTCCAACTGCAACAGTCTGAGCTGAGCTGGTG 468  
Qy 421 AAGCTCGGGGCTCAGTGAAGATTTCCTGCAAGATTCTGGCTAGGCTACGATTCATAGCTCT 480  
Db 469 AAGCTCGGGGCTCCTGAAACTCTCTGTGCAAGCTCTGAGCTCACTTTCACTAGTAT 528  
Qy 481 TGGATGAATGGGTAAGACAGAGCCCTGGACAGGGTCTTGAGTGGATGAGACGGATTAT 540  
Db 529 GCCATGTTCTGGGTTGGCCAGACTCCAGAGAGAGGCTGGAGTGGGTGCGATCCATTAGT 588  
Qy 541 CTTGGAGATGGAAGATTCTAATCTACAAATGGGAATTCGAGGGCAAGGCCATCTACAGCA 600  
Db 589 AGTGG---TGGTTTCACTACTATCCAGACAGTGTGAAGGCCGATTCACCATCTCCAGA 645  
Qy 601 GACAAATCTCCACACAGCCATACATGAGCTCAGCAGCTGAGCCTCTGTGGACTCTGGG 660  
Db 646 GATAATGCCAGGAACATCCTGTATCTGCAATAGCAGTCTGAGTCTGAGGCTGAGGACACGGCC 705  
Qy 661 GTCTATTCTGTGCAAGATCGGGTGTGCTACGTTATGCTATGAGTACTGAGGCTCAGGA 720  
Db 706 ATGATTACTGTGCAAGAGAGAGATAGGGGGTA---CCTCGATGTTGGGGGCGAGGG 762  
Qy 721 ACCTCAGTCACCGTCTCCTCG 741  
Db 763 ACCACGGTCACCGTTTCCTCG 783

RESULT 13  
E35160  
LOCUS E35160 1457 bp DNA linear PAT 31-JAN-2002  
DEFINITION Method for inducing apoptosis.  
ACCESSION E35160  
VERSION E35160.1 GI:18624419  
KEYWORDS JP 2000102389-A/7.  
SOURCE unidentified  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 1457)  
AUTHORS Hagiwara, H., Aotsuka, Y. and Miyahara, J.  
TITLE Method for inducing apoptosis  
JOURNAL Patent: JP 2000102389-A 7 11-APR-2000;  
YOSHIIDE HAGIWARA, HIDEAKI HAGIWARA  
COMMENT OS Artificial Sequence  
PN JP 2000102389-A/7  
PD 11-APR-2000  
PF 29-SEP-1998 JP 1998291441  
PR  
PI HIDEAKI HAGIWARA, YASUYUKI AOTSUKA, JUNICHI MIYAHARA PC  
C12N15/09, A61K31/00, A61K48/00, C07K16/42, C07K19/00, C12N5/10, PC  
C12N15/02//  
PC (C12N5/10, C12R1:91), C12N5/00, C12N5/00, C12N15/00, C12N5/00, PC  
C12R1:91)  
CC  
FH Key Location/Qualifiers  
FT CDS Location/Qualifiers  
1. 1457  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"

ORIGIN  
Query Match 56.5%; Score 418.6; DB 6; Length 1457;  
Best Local Similarity 73.6%; Pred. No. 1.8e-112;  
Matches 554; Conservative 0; Mismatches 184; Indels 15; Gaps 1;  
Qy 1 GATATTCAGATGACAGACTACATCCCTCCTGCTGCTCTCTGGGAGACAGAGTCACC 60  
Db 67 GACATCAGATGAACAGTCTCCATCAGTCTGTCATCCCTTGGAGACACATATCC 126



Search completed: April 26, 2004, 16:55:18  
Job time : 3248 secs